

Pile Sampling White Paper



**State of Idaho
Department of Environmental Quality**

June 2015



*Printed on recycled paper, DEQ June 2015,
PID BIOS, CA code 82993. Costs associated with
this publication are available from the State of Idaho
Department of Environmental Quality in accordance
with Section 60-202, Idaho Code.*

Pile Sampling White Paper

June 2015

**Prepared by
Jeff Myers
AECOM**

This page intentionally left blank for correct double-sided printing.

Table of Contents

Acronyms, Abbreviations, and Symbols	vii
1 Statement of Purpose	1
2 Pile Sampling Challenges	1
3 Quality Systems and Systematic Planning	2
3.1 Data Quality Objectives	3
3.2 Sampling Theory and Practice.....	5
4 Data Quality Objectives.....	6
4.1 Step 1: State the Problem	7
4.2 Step 2: Identify the Goals of the Study	9
4.3 Step 3: Identify Information Inputs	10
4.4 Step 4: Define the Boundaries of the Study	11
4.4.1 Spatial Boundaries	11
4.4.2 Temporal Boundaries.....	14
4.5 Step 5: Develop the Analytical Approach	14
4.6 Step 6: Specify Performance or Acceptance Criteria	16
4.7 Step 7: Develop the Plan for Obtaining Data	20
5 Data Quality Assessment.....	21
5.1 Step 1: Review the DQOs and Sampling Design	21
5.2 Step 2: Conduct Preliminary Data Review	22
5.3 Step 3: Select the Statistical Test.....	22
5.4 Step 4: Verify the Assumptions of the Statistical Test.....	22
5.5 Step 5: Draw Conclusions from the Data	22
6 Incremental Sampling Methodology	22
6.1 Heterogeneity and Its Impact on Sampling and Decision Making	23
6.2 Types of Heterogeneity	25
6.2.1 Constitution Heterogeneity	25
6.2.2 Distribution Heterogeneity	26
6.3 Sampling Correctness	28
6.4 Sampling Errors	29
6.4.1 Fundamental Error	30
6.4.2 Grouping and Segregation Error.....	31
6.4.3 Long-Range and Periodic Heterogeneities and Sampling Errors	32
6.4.4 Increment Delimitation and Extraction Errors	33
6.4.5 Preparation Errors	35
6.5 Applying Incremental Sampling to Piles.....	37
6.5.1 Pile Access Issues	37

6.5.2	Increment Support and Number.....	40
6.5.3	Sampling Tools.....	42
6.5.4	Field Sampling Summary	42
6.5.5	Laboratory Processing and Subsampling.....	43
7	Statistical Evaluation of Sampling Results.....	44
7.1	Incremental Samples	44
7.2	Discrete Samples	45
7.3	Presence or Absence.....	45
8	Process Flow	47
9	Conclusion	48
	References.....	49

List of Figures

Figure 1. Quality system project life cycle for sampling and analysis.	4
Figure 2. Sampling error exceeds laboratory analytical error.....	5
Figure 3. Pierre Gy's (1976) sampling errors.	6
Figure 4. Conceptual site model.	7
Figure 5. Example pile configurations.....	8
Figure 6. Conceptual model of piles.	8
Figure 7. Revised pile conceptual model based on operational factors.....	8
Figure 8. Temperature zone model in a windrow. (a) Cross-sectional view. (b) 3-D cutaway view. Red temperature zone represents the hottest portion of the windrow with internal temperature decreasing with distance from the core.	9
Figure 9. Situation where the decision unit and sampling unit are equivalent.	12
Figure 10. Multiple sampling units in a decision unit.	12
Figure 11. Creating composite samples from discrete field samples (EPA 1995).	13
Figure 12. Composite sampling in a decision unit (EPA 2002a).....	13
Figure 13. Compositing strategies for sampling units within a decision unit (RCRA 2002).	14
Figure 14. Hypothesis testing decision matrix.....	17
Figure 15. VSP input screen.	18
Figure 16. Decision performance goal diagram.....	19
Figure 17. Revised decision performance goal diagram.....	20
Figure 18. Data quality assessment in the quality process.....	21
Figure 19. Conceptual model of heterogeneity and sample support.....	23
Figure 20. Statistical distribution variation based on sample support.	24
Figure 21. Conceptual model of heterogeneity: larger and dispersed contaminants.	25
Figure 22. Constitution heterogeneity.....	26
Figure 23. Heterogeneity in auto fluff.	26
Figure 24. Distribution heterogeneity.	27
Figure 25. Heterogeneity and clustering with two constituents.....	27
Figure 26. Natural clustering with multiple constituents.....	28
Figure 27. Traditional concept of accuracy and precision.	29
Figure 28. Results of correct and incorrect sampling.	29
Figure 29. Nomograph for tracking the fundamental error (Myers 1997).....	31
Figure 30. Field riffle splitter equipment.	32
Figure 31. Laboratory riffle splitters.....	32
Figure 32. Proper (A) and improper (B) use of a riffle splitter.....	32
Figure 33. Thief, auger, and trier sampling devices.	33
Figure 34. Incorrect and correct sampling devices.	33
Figure 35. Increment delimitation and extraction error for spatulas.	34
Figure 36. Increment delimitation and extraction error for scoops.	34
Figure 37. Increment delimitation and extraction error for spatulas or scoops on Japanese slab cakes.....	34
Figure 38. Correct application of the Japanese slab cake technique.....	35
Figure 39. Increment delimitation (left) and extraction (right) in piles.	35
Figure 40. Detailed depiction of Gy's sampling errors.....	36
Figure 41. Flattened pile with incremental sampling grid (ITRC 2012).	38

Figure 42. Cross-sectional view of a windrow composting pile.....	38
Figure 43. Plan view of a windrow composting pile.	39
Figure 44. Increment collection grid for a large, rectangular pile.	39
Figure 45. Increment collection grid for a large, round pile.	40
Figure 46. Core sampling issues.	41
Figure 47. Soil sieving in the field.	41
Figure 48. Flowchart of incremental sampling methodology.	43
Figure 49. Pile sampling process flow.	48

List of Tables

Table 1. Description of systematic planning elements.	3
Table 2. Quality assessment factors.	4
Table 3. Sample size required to demonstrate with at least $(1 - \alpha)\%$ confidence that at least $p\%$ of a lot or batch of waste complies with the applicable standard (no samples exceeding the standard).....	46
Table 4. Sample size required to demonstrate with at least $(1 - \alpha)\%$ confidence that at least $p\%$ of a lot or batch of waste complies with the applicable standard (one sample exceeding standard).	47

Acronyms, Abbreviations, and Symbols

ALARA	as low as reasonably achievable	VSP	Visual Sample Plan
cm	centimeters	μ	micro, one-one millionth (10^{-6})
COC	contaminant of concern		
DEQ	Idaho Department of Environmental Quality		
DQA	data quality assessment		
DQO	data quality objective		
DU	decision unit		
EPA	United States Environmental Protection Agency		
FE	fundamental error		
ISM	incremental sampling methodology		
PAH	polycyclic aromatic hydrocarbon		
PCB	polychlorinated biphenyl		
QA	quality assurance		
QAPP	quality assurance project plan		
QC	quality control		
RCRA	Resource Conservation and Recovery Act		
SER	simple exceedance rule		
STP	sampling theory and practice		
SU	sampling unit		
TCLP	toxic characteristic leaching procedure		
UCL	upper confidence limit		
VOC	volatile organic compound		

This page intentionally left blank for correct double-sided printing.

1 Statement of Purpose

The Idaho Department of Environmental Quality (DEQ) regulates the characterization and disposition of soils, sediments, and other materials that exist in piles and are known or suspected to contain chemical, biological, radiological, nuclear, explosive, or other types of regulated wastes. Whereas the practices and requirements for characterizing and assessing in situ soils and ground water containing contaminants of concern (COCs) are well established, potentially contaminated piles pose unique challenges and require sampling techniques and approaches that exceed routine requirements (Pitard 1993; Myers 1997).

This paper outlines the unique requirements for representative pile sampling and provides technical background for systematic planning, executing, and assessing pile sampling programs. The techniques presented here are intended to assist both regulators and regulated parties in characterizing and disposing pile materials.

2 Pile Sampling Challenges

From a sampling correctness and data quality perspective, sampling cone or ridged piles has been called “a perfect exercise in futility” (Pitard 1993). Nevertheless, for regulatory, economic, or other purposes, piles are sampled routinely. What appears to be a simple exercise harbors unexpected pitfalls that can either penalize a regulated party unnecessarily or exonerate material that will negatively impact human health and the environment. Despite the limitations, applying best practices can recognize and address the challenges so the best possible representation and assessment can be made for regulatory decisions.

By nature, piles differ from in situ material. Anthropogenic processes have created these ex situ masses and have altered the original physical and chemical configurations. These processes can include mixing and/or segregation; exposure to atmospheric accelerants (oxygen, rain, etc.); accelerated and/or modified biological processes; and increased internal temperatures. All of these processes can intentionally or unintentionally alter the nature of the pile constituents as compared to their in situ state. Moreover, the history of the in situ material, pile creation and storage activities, and original contents of the pile may be poorly understood. These post-excavation issues elevate the uncertainty and challenges associated with characterizing piles.

From a physical perspective, piles create access challenges for sampling. Whereas in situ surface topography is typically continuous with gradual slope changes that allow sampling vehicles and equipment to access the subsurface material without extraordinary measures, the abrupt topography and low stability of piles often prohibits standard techniques and equipment from being used. Alternative sampling equipment frequently encounters limitations accessing interior or bottom portions of the pile, where activities related to biological activity, moisture accumulation, and other factors could create heterogeneous areas harboring COCs.

This paper addresses the unique challenges associated with obtaining representative samples from the naturally heterogeneous soil materials in piles along with the zonal heterogeneities, which are both compounded by access and sampling equipment limitations. Recognizing that representativeness is an elusive property that captures correctness, accuracy, and precision, the

intent of this paper is to identify known types of challenges and to provide scientific and engineering approaches to meeting these challenges. Because decisions made concerning the disposition of a pile may have adverse consequences on human health and the environment, this paper will address the technical issues in the context of the US Environmental Protection Agency's (EPA's) quality assurance (QA) and quality control (QC) measures, which are intended to mitigate the uncertainty associated with pile characterization.

3 Quality Systems and Systematic Planning

This section provides an overview of the systematic planning process necessary for pile sampling. It discusses where pile sampling fits into the graded approach used by EPA (2001a,b), noting the special quality challenges posed by pile sampling. EPA uses quality systems to ensure the acceptability of products, services, or information. Quality systems typically integrate multiple elements including management, technical, and administrative (policies and objectives, procedures and practices, organizational authority, roles and responsibilities, and accountability). With quality defined as either "conformance to requirements" or "fitness for use," quality systems ensure the desired level of conformance or fitness is achieved.

Systematic planning is a fundamental component of environmental quality systems. Quality systems structure processes to describe the policies and procedures necessary to ensure work processes, products, or services satisfy expectations or specifications. For environmental applications, data collection through sampling and analysis must conform to quality guidelines. Without proper planning prior to data collection, data will likely fail to meet specific project needs. The quality system provides a blueprint for obtaining project-specific data of appropriate quantity and quality. Experience has shown the benefits of systematic planning before data collection begins. Without it, data quality suffers and valuable resources may be wasted.

EPA's systematic planning processes follow the time-tested scientific method to ensure an objective approach and acceptable results. The scientific method involves formulating a question, initiating background research, constructing a hypothesis, testing the hypothesis, analyzing data, drawing a conclusion, and reporting results. The results address specifically whether the hypothesis was correct. Adapting the scientific method to environmental characterization and sampling, systematic planning guides a project's logical development, plans for efficient use of scarce resources, and provides transparency of intent and direction, soundness of project conclusions, and appropriate levels of documentation for peer and regulatory review. Table 1 lists key systematic planning elements (EPA 2006a), which form the basis of the data quality objectives (DQO) process.

Table 1. Description of systematic planning elements.

Element	Description
Organization	Identify and involve the project manager, sponsoring organization and responsible official, project personnel, stakeholders (e.g., all customers and suppliers), scientific experts, etc.
Project Goal	Describe the project goal, objectives, and study questions and issues.
Schedule	Identify project schedule, resources (including budget), milestones, and any applicable requirements (e.g., regulatory requirements, contractual requirements).
Data Needs	Identify the type of data needed and how the data will be used to support the project's objectives.
Criteria	Determine the quantity of data needed and specify performance criteria for measuring quality.
Data Collection	Describe how and where the data will be obtained (including existing data) and identify any constraints on data collection.
Quality Assurance	Specify needed QA and QC activities to assess the quality performance criteria (e.g., QC samples for both field and laboratory, audits, technical assessments, performance evaluations, etc.).
Analysis	Describe how the acquired data will be analyzed (either in the field or the laboratory), evaluated (i.e., QA review/verification/validation), and assessed against their intended use and the quality performance criteria.

3.1 Data Quality Objectives

The DQO process provides the core systematic planning process for environmental characterization. Its relationship to quality systems and pile sampling is introduced here.

EPA's quality system for sampling and analysis programs consists of three phases (Figure 1): planning, implementation, and assessment. The first phase includes the DQO process, a seven-step systematic planning process that develops a defensible quality assurance project plan (QAPP) or waste analysis plan. The DQO process enables stakeholders to ask the right questions and develop appropriate answers and procedures to address specific project needs so appropriate and representative data are collected during the implementation phase and are acceptable in the assessment phase. See section 4 for a detailed discussion on the DQO process.

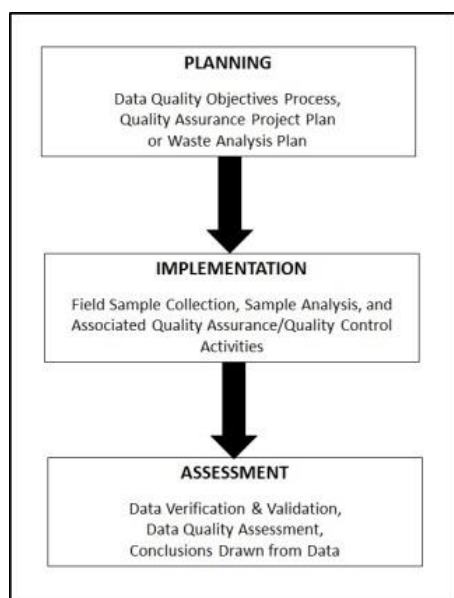


Figure 1. Quality system project life cycle for sampling and analysis.

The DQOs guide the field sampling program and analytical laboratory efforts under the QAPP or waste analysis plan, which also emanate from the DQO process and make up the implementation phase. These documents detail the number of samples necessary to meet decision-making goals, where and how the samples are to be taken, the methods by which the samples will be analyzed, the necessary laboratory QA/QC protocols required, and the levels of confidence associated with decisions.

Finally, the assessment phase involves evaluating the data collected to confirm they meet the objectives and standards prescribed in the DQOs. The quality of decisions made for waste remediation, treatment, or disposal relies on the quality of the data used to make the decisions. The original objectives cannot be met if inferior data are collected. Table 2 presents EPA's general assessment factors for data quality. See section 5 for a detailed discussion on data quality assessment.

Table 2. Quality assessment factors.

Factor	Description
Soundness	The extent to which the scientific and technical procedures, measures, methods, or models employed to generate the information are reasonable for and consistent with the intended application.
Applicability and Utility	The extent to which the information is relevant for the agency's intended use.
Clarity and Completeness	The degree of clarity and completeness with which the data, assumptions, methods, quality assurance, sponsoring organizations, and analyses employed to generate the information are documented.
Uncertainty and Variability	The extent to which the variability and uncertainty (quantitative and qualitative) in the information or the procedures, measures, methods, or models are evaluated and characterized.
Evaluation and Review	The extent of independent verification, validation, and peer review of the information or of the procedures, measures, methods, or models.

3.2 Sampling Theory and Practice

Analytical laboratories have traditionally shouldered blame for sample result anomalies. The development of sampling theory and practice (STP) in the mining industry allowed important technology to be transferred to environmental applications, which assists in separating laboratory error from field sampling error. A brief history appears below.

Superficially, sampling soils and other heterogeneous materials appears to be a straightforward and routine exercise. However, this expectation frequently collides with reality to produce confusing and unpleasant results—unexplained variability that necessitates updates to the conceptual models, replicate laboratory results that differ markedly and complicate decision making, and the inability to achieve statistical quality goals due to insufficient numbers of samples. All of these concerns can disrupt regulatory acceptance processes and initiate repeat sampling programs to achieve acceptable results.

STP was developed by Pierre Gy (1976) to address the challenges encountered in sampling particulate materials. STP is especially useful for sampling programs concerned with trace constituents (under about 1% or 10,000 parts per million), which are prevalent in environmental characterizations. Gy recognized how the inherent soil heterogeneities impact sample results, making them highly inaccurate in most cases. Until this point in time, industry had misattributed sampling errors as laboratory errors.

Gy's revelation provided a paradigm shift in thinking and practical tools to minimize sampling error. Instead of instinctively attributing differing replicate results to laboratory error, the focus shifted to addressing sampling error. Figure 2 presents an example of this phenomenon. The figure shows sample results obtained using field analytical and fixed laboratory analysis for material obtained at a single location. This figure also applies to situations where field duplicates are analyzed by a laboratory or by different laboratories.

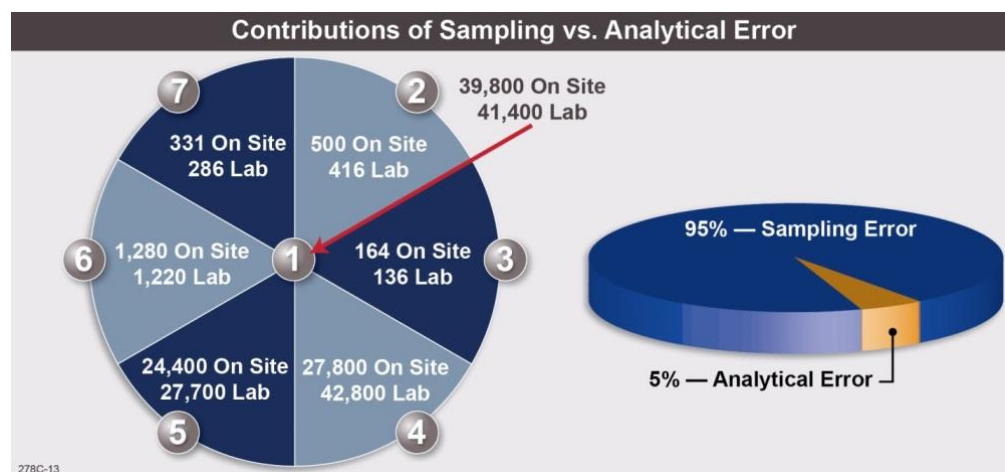


Figure 2. Sampling error exceeds laboratory analytical error.

Figure 2 demonstrates the validity and importance of Gy's new paradigm—laboratory (analytical) errors are very low compared to sampling errors. According to studies by EPA and others, sampling errors contribute on average 95% of the variability and error, whereas laboratory errors contribute only 5%. Given this insight, the need to minimize field sampling

errors becomes obvious. Using Gy's methods for sampling soils is recommended by EPA and others (EPA 2002a; Pitard 1993; Myers 1997; ITRC 2012).

Gy identified, categorized, and organized the seven sampling errors into a unified practice (Figure 3). Figure 3 shows how error is associated with different aspects of the sampling process, starting with the inherent heterogeneity of the material and ending with preparation error. The DQO process is used to plan a field sampling program that will minimize each of the seven sampling errors. See section 6 for a detailed discussion on STP.

Gy's Seven Sampling Errors	
Discrete Model	
Fundamental Error	
Incurred due to Constitutional Heterogeneity (particle size distribution)	
Grouping & Segregation Error	
Results from Distributional Heterogeneity (clustering)	
Increment Delimitation Error	
Incorrect sampling equipment	
Increment Extraction Error	
Incorrect material collection	
Continuous Model	
Long-Range Heterogeneity	
Non-random spatial distribution	
Periodic Heterogeneity	
Cyclic temporal and/or spatial distribution	
Human Factors	
Preparation Error	
Loss, contamination, or alteration of sample/subsample	

Figure 3. Pierre Gy's (1976) sampling errors.

4 Data Quality Objectives

This section provides an overview of EPA's seven-step DQO process, with special emphasis on pile sampling application. EPA 2006a is referenced as the source of detailed process descriptions; this section is not intended to replicate DQO guidance in total.

DQOs can be broadly summarized by three questions (Myers 1997):

1. What is the question to be answered?
2. What is the statistical population of interest?
3. What is the acceptable level uncertainty?

The three questions address much of what is covered in DQO Steps 1 through 6. Summary descriptions of the seven DQO steps appear below. Detailed DQO guidance can be found in EPA 2006a.

4.1 Step 1: State the Problem

This discussion covers the applicable Step 1 activities listed below, with special emphasis on activities 1 and 3, as these vary compared to traditional environmental sampling:

1. Develop a concise description of the problem.
2. Identify the leader and members of the planning team.
3. Develop a conceptual site model of the pile.
4. Determine resources—budget, personnel, and schedule.

Activity one, problem definition, is a crucial step in successful characterization. Without a clear understanding of the problem, plans and work efforts will likely not address the goals and objectives; thus, unfocused data collection will produce insufficient or inappropriate sample data for decision-making. Project success hinges on a clearly defined and simply stated problem.

Closely linked to the problem statement is activity three, developing a conceptual model of the environmental hazard. Environmental problems are often complex; simplification can be achieved by portraying the conceptual model as a diagram showing physical, chemical, media, and operational aspects of the project. Because the conceptual model will serve as the basis for project inputs and decisions, project success depends on developing an accurate conceptual model. Figure 4 shows a conceptual site model for a typical environmental project.

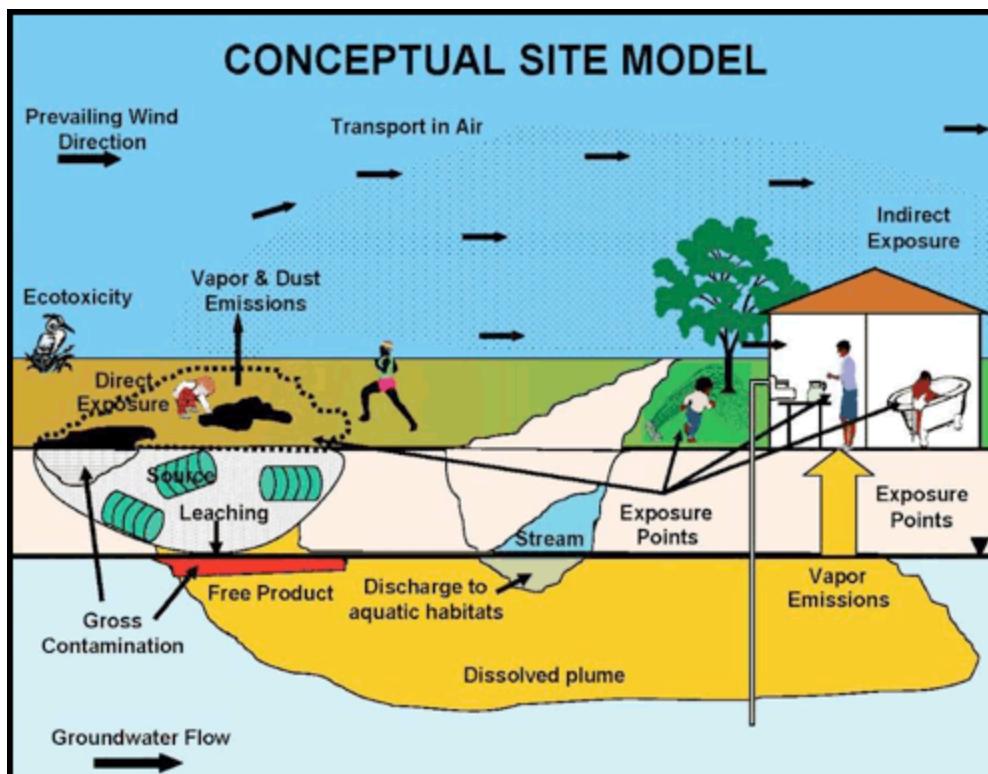


Figure 4. Conceptual site model.

Piles are infinite in their sizes, shapes, and composition but tend to have simpler conceptual models. Pile examples appear in Figure 5 and exhibit the aforementioned characteristics: variability in size, shape, and composition.



Figure 5. Example pile configurations.

Figure 6 is a conceptual model of piles that have potentially different characteristics based on their colors, with the colors suggesting different origins and compositions. In this situation, it may be appropriate to sample each pile separately.

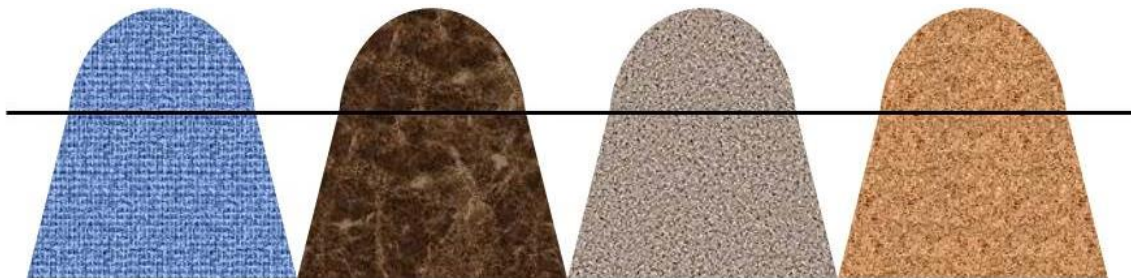


Figure 6. Conceptual model of piles.

Because the pile shapes create potential logistical issues for sampling access, the horizontal line proposes an elevation for flattening the piles. Figure 7 shows the revised conceptual pile configuration after a flattening process. Note how the four mounded shapes are replaced with three rhombohedral shapes and one trapezoidal shape. Note further how a vertical core sample on one side of a central rhombohedron will intersect two types of material, causing a sampling bias by mixing materials. Thus, by flattening the piles, the conceptual model for sampling changes. This implies the pile conceptualization task may be a process subject to revision, rather than a one-time exercise.

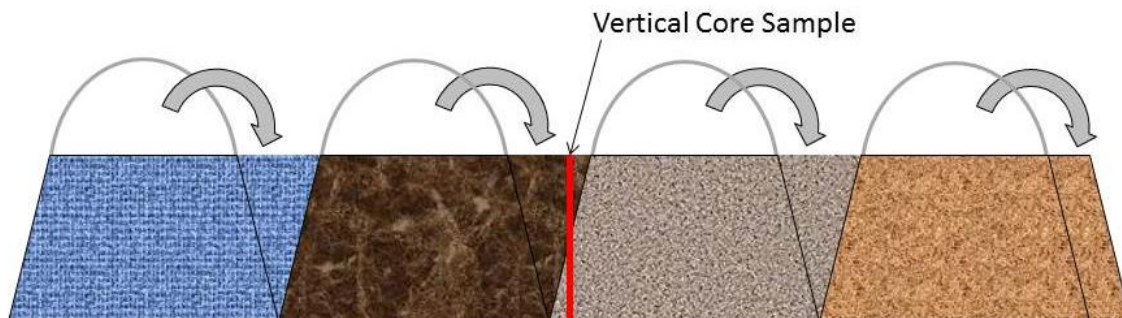


Figure 7. Revised pile conceptual model based on operational factors.

Figure 8 is a conceptual model of the temperature zones in a windrow composting scenario, showing increasing temperatures as one moves from the pile exterior into the interior. Increased heat in a zone may influence the rate of constituent degradation and elimination. Actual field temperature measurements may be required to validate the conceptual model, which might be used to estimate the time required for COCs to degrade below regulatory levels.



Figure 8. Temperature zone model in a windrow. (a) Cross-sectional view. (b) 3-D cutaway view. Red temperature zone represents the hottest portion of the windrow with internal temperature decreasing with distance from the core.

Conceptual models should capture the unique aspects and characteristics pertinent to a specific project. Conceptual model development may require multiple attempts, with experiments or testing to acquire information to refine it.

Step 1 is the starting point for the remaining DQO activities by defining what the problem is, who will help solve it, and how the parties will conceptualize the problem. Detailed descriptions of Step 1 activities are found in EPA 2006a (QA/G-4), pages 15–20.

4.2 Step 2: Identify the Goals of the Study

Organizing multiple decisions can be complicated when objectives include sampling and analysis of multiple COCs for decision-making. The goal of this section is to provide guidance when the need for multiple COCs complicates the necessary sampling approaches.

DQO Step 2 identifies key questions the study needs to address along with the potential actions associated with sample outcomes that answer the questions. Step 2 is the opportunity to organize multiple questions and/or decisions that may result from the sampling. The principle activities of Step 2 include the following:

1. Identify the principal study question(s).
2. Consider alternative outcomes or actions that can occur.
3. For decision problems, develop decision statement(s) and organize multiple decisions.
4. For estimation problems, state what needs to be estimated along with key assumptions.

The principal study question is formulated based on the problem statement developed in Step 1. Problem statements fall into two categories, those that initiate a fixed decision and those that produce an estimation. Decisions require choosing a course of action, usually from two

alternatives. Estimations address the magnitude of an environmental parameter or characteristic, perhaps to be used in a modeling effort.

Initially, a single study question should be established, which can be amplified later in the DQO process to include related issues and questions. Examples of principal study questions include the following:

Decision Problems

- Do toxic characteristic leaching procedure (TCLP) for metals results exceed acceptable limits?
- Do total metals results exceed the 20X rule, requiring TCLP?
- Do volatile organic compound (VOC) and/or polycyclic aromatic hydrocarbon (PAH) results exceed criteria for landfill closure?
- Do VOC and PAH results exceed criteria for using soils as backfill material?
- Do total petroleum hydrocarbon results exceed criteria for certain landfill disposal?
- Are fecal coliform concentrations below regulatory limits?

Estimation Problems

- What is the decay curve for explosives during windrow composting?
- What type/quantity of amendments optimize septage degradation?
- How much ammonia will be released during explosives composting?

At the conclusion of Step 2, the following outputs should be present:

- A well-defined principal study question
- A compilation of alternative outcomes or actions resulting from efforts to resolve the principal study questions
- For decision problems, a list of decision statements that address the study question
- For estimation problems, a list of estimation statements that address the study question

Detailed descriptions of Step 2 activities are found in EPA 2006a (QA/G-4), pages 21–26.

4.3 Step 3: Identify Information Inputs

DQO process Step 3 addresses the types and sources of information necessary for decision statement or estimation resolution. Whereas sampling and analysis work is presumed, this phase confirms whether these activities are necessary because existing data are occasionally sufficient for decision making. This step determines the information types needed for sampling and analytical methods necessary and whether the available techniques will meet acceptance criteria. The main activities in this step include the following:

1. Identify types and sources of information needed for decisions or estimates.
2. Identify the basis of information to guide or support choices in the subsequent DQO steps.
3. Select appropriate sampling and analysis methods.

To identify the types of information needed, the following types of questions are applicable:

- Is information on the physical properties of the media required?

- Is information regarding the chemical characteristics of the matrix needed?
- Can existing data be used to make the decision or produce the estimate?
- Do we need to collect new measurements on environmental characteristics?

Step 3 activities are forward-looking to the activities performed in Step 5, where the sampling and analytical approaches are developed. The planning team should compile lists of the physical sampling and laboratory analytical methods capable of meeting projects goals. If the list fails to include appropriate methods, the team should return to Step 2 and restructure the study goals.

The outputs from DQO Step 3 are as follows:

- Lists of environmental characteristics that will resolve the decision or estimate and potential sources for the desired information inputs
- Information on the number of variables that will need to be collected
- The type of information needed to meet performance or acceptance criteria
- Information on the performance of appropriate sampling and analysis methods

Detailed descriptions of Step 3 activities are found in EPA 2006a (QA/G-4), pages 27–30.

4.4 Step 4: Define the Boundaries of the Study

DQO Step 4 imparts spatial and temporal context to individual samples and groups of samples. This is a key area and is related to the conceptual model of the pile and the conceptual model of heterogeneity. The relationship to establishing sampling units (SUs) and decision units (DUs) is presented in detail in section 4.4.1.

Step 4 focuses on the following activities:

1. Define the target population and its relevant spatial boundaries.
2. Define what constitutes a sampling unit.
3. Specify temporal boundaries and other practical constraints.
4. Specify the smallest unit on which decisions will be made.

4.4.1 Spatial Boundaries

The target population for pile sampling is the pile or piles of concern. However, the conceptual model may identify zones of differing physical, chemical, biological, temperature, stratification, or other attributes that need to be characterized individually. These zonal or area boundaries must be identified clearly for both spatial and temporal considerations.

For reasons related to the conceptual model, physical constraints, convenience, and decision-making considerations, SUs must be defined. Simultaneously, DUs must also be defined. The SU and DU concepts are closely related, with the distinction between the two as follows:

- A **DU** is the smallest volume of soil or other material on which a *decision* will be made as a result of sampling.
- An **SU** is a volume of soil or other material from which *samples* (discrete, composite, or incremental) will be collected to characterize the volume.

Figure 9 shows the simplest case where the DU is the SU and vice versa. Figure 10 shows two possible SU configurations within a DU.

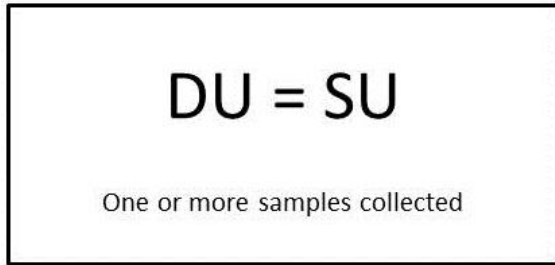


Figure 9. Situation where the decision unit and sampling unit are equivalent.

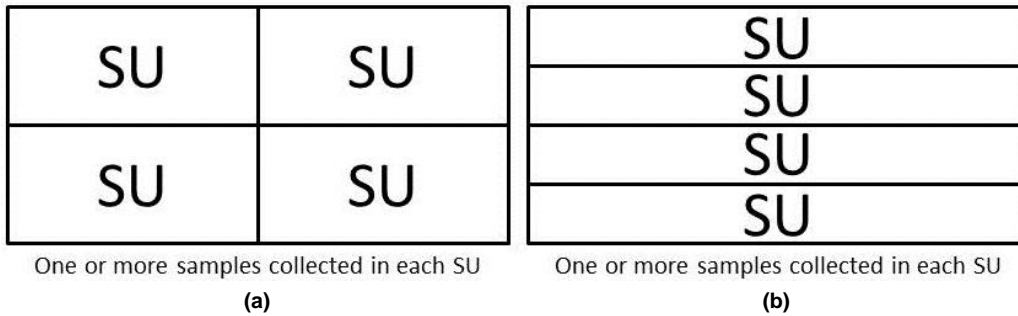


Figure 10. Multiple sampling units in a decision unit.

DUs may be established based on multiple considerations, including the following:

- Risk and the scale of decision making (size/volume of DU)
- Technological considerations (most efficient size/shape for sampling, composting, or other process)
- Temporal considerations (exposures [chemical, radiation/ALARA], treatment processing times)
- Financial scale (processing costs)
- Other (e.g., specialized verification sampling)

SUs and DUs may be sampled using a variety of methods and categories. *Discrete samples* (also known as grab samples) and *composite samples* (also known as integrated samples) differ with regard to how the physical material is collected in the field before laboratory analysis. Statistical and judgmental approaches are categories that refer to how the sampling patterns or sample locations were established.

Discrete samples are collected from material at a single location, followed by analytical laboratory analysis of a portion of the sample material. Discrete samples may be collected randomly, on stratified random grids, on fixed grids of various configurations, by judgmental selection, or some combination of these approaches. Fixed grids may be square, rectangular, triangular/hexagonal, radial, cluster, or another form (Myers 1997). Randomizing the grid origin is recommended to reduce potential sampling bias. Discrete samples are often used when the desired information is the type and shape of the statistical distribution.

Composite samples are the result of combining sample material from multiple locations or time intervals (several discrete samples) to create a single sample for laboratory analysis. Composite samples are often used when (1) the mean (average) of a constituent in a material (air, water, soil, etc.) is desired and (2) when laboratory costs need to be minimized. Figure 11 shows the basic concept of combining field samples into composite samples.

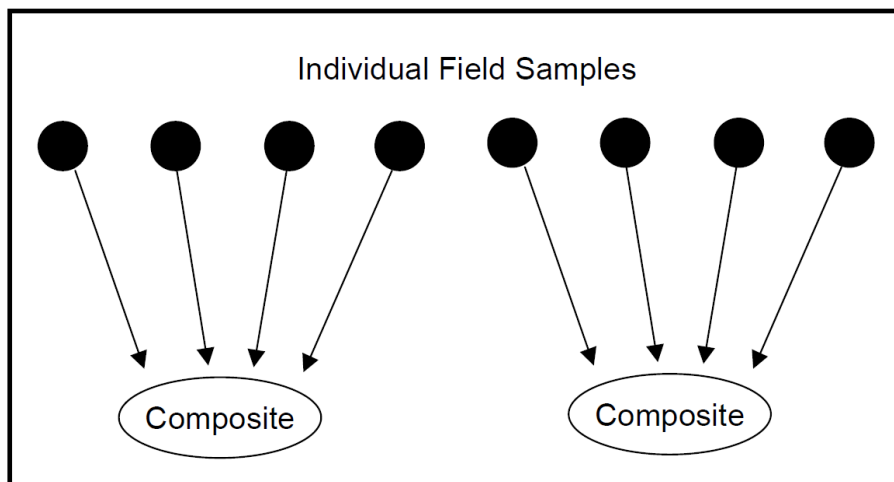


Figure 11. Creating composite samples from discrete field samples (EPA 1995).

Figure 11 through Figure 13 show an assortment of field collection approaches to composite sampling. Figure 12 shows samples representing three random sample locations in a DU being combined into one composite sample, with this process repeated three times to produce a total of three composite samples from the nine field sample locations. Figure 13 presents two types of composite designs: (a) where one discrete sample is collected from each of the six SUs using a fixed square grid, with the six discrete samples to be combined into a single sample, and (b) where four discrete samples are collected from each SU and combined into a composite sample. Scenario “a” produces four composite samples that are analyzed at the laboratory and the results averaged to produce an estimate of the DU mean concentration. Scenario “b” produces six composite samples that are analyzed and averaged.

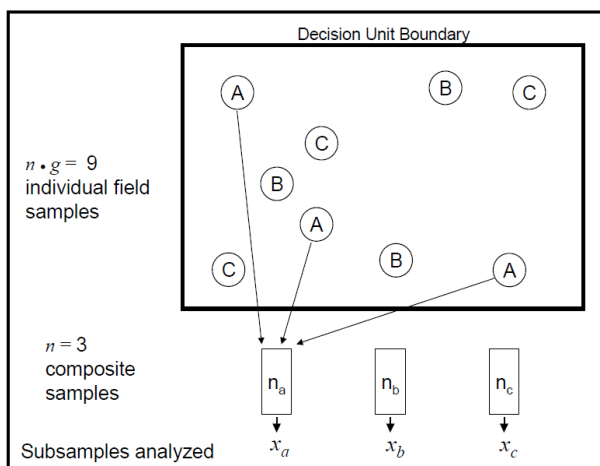


Figure 12. Composite sampling in a decision unit (EPA 2002a).

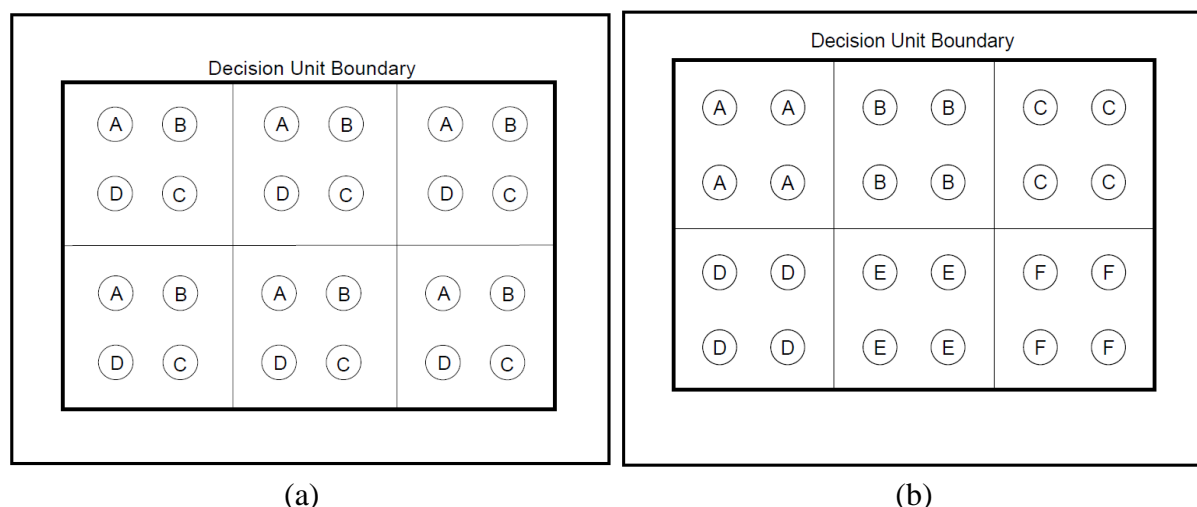


Figure 13. Compositing strategies for sampling units within a decision unit (RCRA 2002).

Incremental sampling is an advanced form of composite sampling that addresses the sampling errors shown in Figure 3. Incremental sampling is discussed in detail in section 6.

4.4.2 Temporal Boundaries

Piles subject to composting or other time-related processes require DQOs to understand how pile conditions vary over the course of treatment and the related results achieved or actions necessary. For example, to establish project costs, the time necessary to compost a batch windrow material must be known or estimated, especially if multiple processing batches are planned. Similarly, the amount of undesirable gases (e.g., ammonia) emitted and when minimum and maximum off-gassing periods occur must also be known. If the project is large, construction time or delays as well as contractor availability and contingencies should also be considered.

When the spatial and temporal boundaries have been established, they should be documented along with the assumptions and developmental logic. The outputs from Step 4 are as follows (EPA 2006a):

- Definition of the target population with detailed descriptions of geographic limits (spatial boundaries)
- Detailed descriptions of what constitutes an SU
- Time frame appropriate for collecting data and making the decision or estimate, together with those practical constraints that may interfere with data collection
- The appropriate scale for decision making (DU)

Detailed descriptions of Step 4 activities are found in EPA 2006a (QA/G-4), pages 31–37.

4.5 Step 5: Develop the Analytical Approach

The statistical framework and decision actions begin to take form in DQO Step 5. Step 5 focuses on developing an analytical approach to guide the interpretation of study results and conclusions. The term *analytical approach* refers to the data and/or statistical analysis applied to the sample data collected during the study. Analytical approach in this context does not refer to the

laboratory method or analytical process used to obtain contaminant concentrations. Laboratory methods, detection limits, and related requirements are determined in DQO Step 3.

The outputs developed in Steps 1–4 need to be integrated with the statistical parameter selected and validated for consistency toward achieving the stated goals. Unambiguous “if...then...else...” statements should be developed for decision problems; clear specification of the estimator (statistical function) to be used is necessary for estimation problems.

Below are the principal activities for Step 5:

1. Specify appropriate population parameters for decisions or estimates.
2. For decision problems, choose an appropriate action level(s) and generate an “if...then...else...” decision rule related to the chosen statistical parameter.
3. For estimation problems, specify the estimator (statistical function) and the estimation procedure.

Any valid and DQO-appropriate statistical parameter may be chosen for the decision rule. Examples of common selections include the mean (arithmetic or geometric), median, or percentile. Other parameters such as presence/absence through the application of the simple exceedance rule (SER) (EPA 2002a) or hypergeometric distribution may also be appropriate to meet DQO goals. The mean is the most commonly selected parameter and is highly applicable to pile sampling. The mean is frequently adjusted to provide an upper confidence limit (UCL) on the mean to account for variability and uncertainty in the sampling and analysis. The level of confidence chosen should be identified in the DQO process.

The mean or other parameter can then be compared to an appropriate regulatory action level under the “if... then... else...” decision rule. For example, a decision statement could take the form of:

“If the true arsenic concentration in the soil pile exceeds x ppm, then the pile will be disposed as hazardous waste.”

Because soil, rock, and heterogeneous soil-like materials typically compose pile material, incremental sampling is used frequently to minimize sampling error. By design, incremental sampling produces an estimate of the mean of the parameter of concern. As such, pairing the use of the mean as the statistical parameter with the incremental sampling approach is both appropriate and applicable to pile characterization. In contrast, if the median or percentile is desired as the statistical parameter, a discrete sampling method may be more appropriate than an incremental sampling approach because the median relates to a discrete value in a data distribution. Incremental sampling is discussed in detail in section 6.

The outputs of DQO Step 5 are as follows (EPA 2006a):

- Identification of the population parameter(s) most relevant for making inferences and conclusions on the target population
- For decision problems, the “if... then... else...” theoretical decision rule based on a chosen action level
- For estimation problems, the specification of the estimator to be used

Detailed descriptions of Step 5 activities are found in EPA 2006a (QA/G-4), pages 39–44.

4.6 Step 6: Specify Performance or Acceptance Criteria

Managing decision performance is a crucial function of the DQO process. The issues addressed in this step are similar to those in DQO Step 5.

Despite the best efforts and best practices implemented in the DQO, sampling, and analytical processes, Step 6 deals with the reality of imperfect data and the uncertainty associated with subsequent decision making. When comparing the true value of the selected statistical parameter to the action level, the existence of uncertainty introduces the possibility of an incorrect decision, which could cause nonhazardous material to be declared as hazardous or vice versa. In the first case, unnecessary expense is incurred; in the second case, inappropriate risks are introduced to human health and the environment. Step 6 is where boundaries are set for decision errors considered tolerable.

The following actions are taken during Step 6:

- For decision problems:
 1. Specify the decision rule as a statistical hypothesis test(s).
 2. Examine the consequences of making incorrect decisions.
 3. Place acceptable limits on decision errors.
- For estimation problems:
 1. Examine and evaluate the consequences of incorrect decisions.

Statistical hypothesis tests are the generally accepted method for making environmental and regulatory decisions on sample data. The hypothesis test decision is based on whether sufficient evidence exists to accept a null hypothesis (H_o) (assumed true state of nature or baseline condition), or whether to reject the null hypothesis and accept the alternative hypothesis (H_A) (EPA 2006a; Myers 1997). Example hypothesis statements follow:

H_o : *The average arsenic concentration in the pile is less than the action level.*

H_A : *The average arsenic concentration in the pile is equal to or greater than the action level.*

Figure 14 presents a matrix of decision errors and correct decisions for hypothesis tests. The matrix works in the following manner: if one hypothesizes the baseline condition to be true and it is, a correct decision is made. Similarly, if the alternative condition is true and the alternative condition is chosen, a correct decision is again made. Note that two of the four possible decisions are correct while two are incorrect. Decision errors occur in the following situations. If the baseline condition is assumed to be true but is actually false, then a false acceptance decision error is made. Conversely, if you decide the alternative condition is true but it is actually false, a false rejection is made.

Decision You Make by Applying the Statistical Hypothesis Test to the Collected Data	True Condition (Reality)	
	Baseline Condition is True	Alternative Condition is True
Decide that the Baseline Condition is True	Correct Decision	<i>Decision Error (False Acceptance)</i>
Decide that the Alternative Condition is True	<i>Decision Error (False Rejection)</i>	Correct Decision

Figure 14. Hypothesis testing decision matrix.

A hypothesis test parallels closely the process used in our legal system, where an accused person is assumed to be innocent (null hypothesis H_0). If the jury decides the accused person is guilty but the person is innocent, a false rejection of innocence occurs and the person is wrongly convicted. Conversely, if the person is guilty but the jury decides the person is innocent, then a false acceptance of innocence occurs. In this case, a criminal is set free and remains a continued risk to the public.

False rejections are termed Type I errors; false acceptances are termed Type II errors. The probability of a Type I error is called alpha (α) and represents the hypothesis test's level of significance. An alpha of 0.05 yields a $1 - \alpha$ probability, or a 95% confidence, in the test. The probability of a Type II error is called beta (β) and refers to the hypothesis test's power. Calculating $1 - \beta$ provides the power of the test.

Type I and II errors are generally managed in tandem, although Type I errors may be managed alone without considering Type II errors. Typical environmental hypothesis statements call for more stringent alpha values (0.10 to 0.01), whereas beta values are commonly around 0.20. More stringent alpha values are more protective of human health and the environment; however, regulated parties may opt to select more stringent beta values to minimize the risk of falsely declaring a waste lot to be hazardous. Note also that more stringent alpha or beta values increase the number of samples needed to prove the hypothesis at the desired levels of confidence, thus increasing the sampling and analytical costs.

The error tolerances are negotiated during the DQO process. Common starting points are 5% alpha error/significance (95% confidence) and 20% beta error (80% power). Depending on the impact severity of a decision error, the error tolerance may be adjusted up or down as long as the parties agree.

Public domain computer software is available to manage Type I and II decision errors and to determine appropriate levels of sampling. One such program is Visual Sample Plan (VSP), developed by the US Department of Energy. Figure 15 shows a set-up page for VSP. This particular screen shows a commonly used approach for piles and other discrete DUs: comparing an average contaminant concentration to a fixed threshold. Numerous types of set-up pages are available in VSP depending on the type of statistical analysis needed.

The VSP program requires several inputs. Input parameters allow users to customize the analysis and include fundamental assumptions, desired levels of confidence and significance (Type I and Type II error limits/alpha and beta), the population standard deviation, the regulatory action level, and the width of the gray region (explained below), which is based on the average concentration. Based on the input parameters and assumptions, the program calculates the

minimum number of samples necessary to fulfill the input requirements, which equals for the parameters shown in the figure.

True Average vs. Fixed Threshold

Average vs. Fixed Threshold | Sample Placement | Costs | Data Analysis | Analytes

I **can** assume the data will be normally distributed. For Help, highlight an item and press F1

I want to use **ordinary** sampling.

These design parameters apply **Analyte 1**

Specify Null Hypothesis:

I want to assume the site **unacceptable (dirty)** until proven otherwise.
(Assume the true mean \geq action level.)

Specify False Rejection Rate (alpha) and Action

I want at least **95.0** % confidence that I will conclude the site is unacceptable (dirty) if the true mean is at or above the action level of **10** units.

Specify Width of Gray Region (delta) and False Acceptance Rate (beta):

If the true mean is **1** units below the action level (that is, 9 units)
then I want no more than **20.0** % chance of incorrectly accepting the hypothesis that the site is unacceptable (true mean \geq action level).

The estimated standard deviation due to sampling and analytical variability is **MQQ**
2 units.

☐ I want to include historical (existing) samples in this design

Minimum Number of Samples for Analyte 1: 27

Minimum Number of Samples in Survey Unit: 27

Figure 15. VSP input screen.

Figure 16 presents an example VSP output graph called a decision performance goal diagram. The graph provides a visual representation of the analysis and documents the number of samples necessary to determine if the average concentration can be considered below an action level at the specified level of confidence and power.

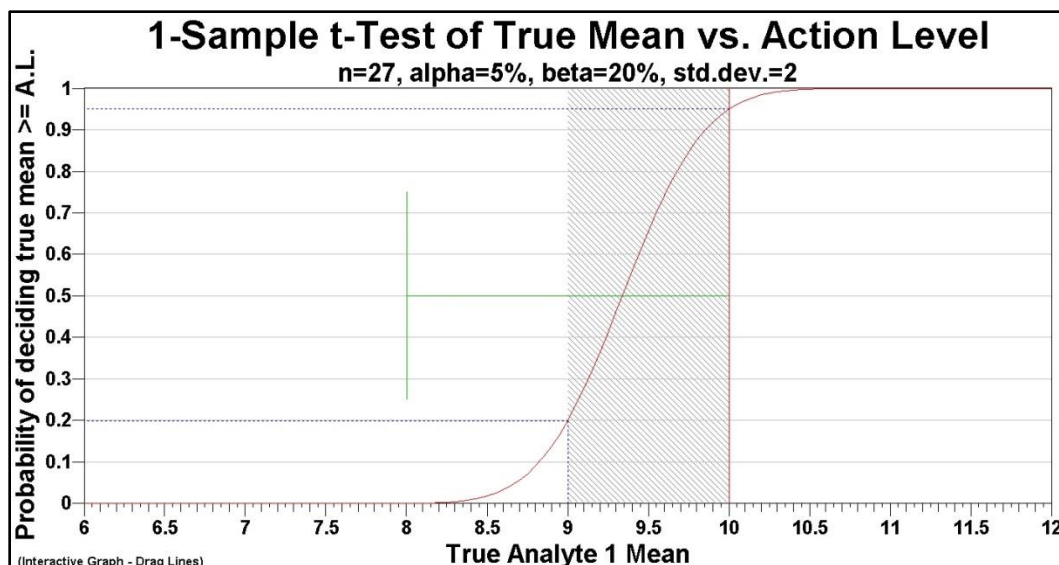


Figure 16. Decision performance goal diagram.

The decision performance goal diagram echoes the component features of the input screen. The x -axis plots the concentration or value of the parameter being evaluated. The y -axis plots the probability of deciding the true parameter mean is equal to or exceeds the action level. The green, horizontal “T” plots the size of the standard deviation of the parameter of interest. The shaded area of the graph is called the “gray region.” This is a zone of uncertainty where insufficient sample data exist to make decisions with the desired levels of confidence. The blue, horizontal dashed lines plot the desired levels of statistical confidence. In the lower left portion of the graph, the horizontal blue line represents beta; in the upper left portion, the blue line plots alpha. The vertical red line on the right side of the graph is the action level for the parameter of interest. The red line starting in the lower left corner and ending as a flat line at the top of the graph after passing through the gray region is the performance curve.

When the line enters the gray region at the lower left portion and starts to rise, the decision error rate is only marginally unacceptable. When the performance curve line reaches the 0.5 level on the y -axis, there is a 50/50 chance of making a wrong decision, the same odds as tossing a coin. When the performance curve approaches the red vertical action level line, the waste is very likely to be hazardous but has not reached the necessary confidence level for decision making. It is only when the performance curve is outside the gray region that sufficient evidence (sampling) exists to make the appropriate risk-quantified decision.

The width of the gray region influences the number of samples necessary. If the anticipated sample mean is close to the action level, the decision is more difficult and more data are required. Figure 17 shows an analysis where the expected mean has been changed from 9.0 to 9.5. In this situation, 101 samples are needed instead of the previous result of 27 samples.

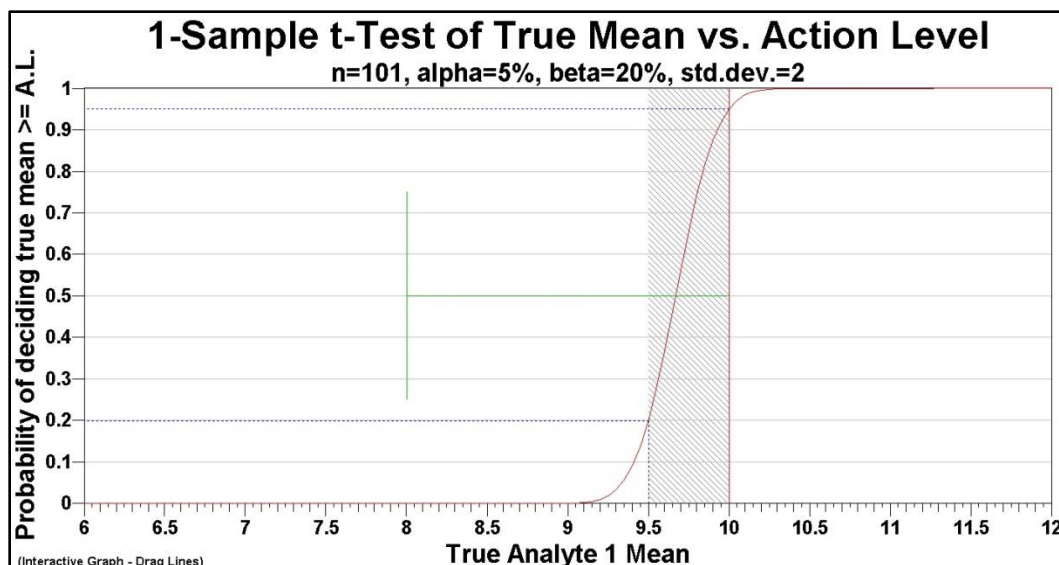


Figure 17. Revised decision performance goal diagram.

VSP information and download links are available at <http://vsp.pnnl.gov/>. The VSP manual can be accessed at <http://vsp.pnnl.gov/docs/PNNL%2019915.pdf>. Additional information on hypothesis testing and the gray region is provided in EPA’s DQO document (EPA 2006a).

Detailed descriptions of Step 6 activities are found in EPA 2006a (QA/G-4), pages 45–70.

4.7 Step 7: Develop the Plan for Obtaining Data

The final DQO step is to synthesize a coherent and compliant plan with the requirements set forth in steps 1–6. Step 7 integrates the “homework” performed in Steps 1–6 to create a final sampling design for the project. Decisions regarding the type of sampling to be performed (discrete, composite, incremental, etc.), how to achieve representativeness of the material being sampled, the types of sampling tools, whether statistical or judgmental sampling applies best, the laboratory analytical methods and detection limits required, cost optimization, and other factors must be finalized and documented. The activities in Step 7 are as follows:

1. Complete all information and outputs generated in Steps 1 through 6.
2. Use information to identify alternative sampling and analysis designs.
3. Select and document a design that will yield data to best achieve desired performance or acceptance.

EPA provides guidance regarding developing a sampling and analysis plan (EPA 2002b). The following types of information are needed to create an effective sampling design (EPA 2006a):

1. The objectives and intended use of the data (e.g., statistical hypothesis testing, estimation)
2. DQO process outputs from Steps 1–6 (the conceptual model, variables of interest, spatial and temporal boundaries, performance or acceptance criteria for the data collected)
3. Background information (site properties, technical characteristics of the contaminants and media, regulatory requirements, known spatial/temporal contamination patterns)

The outputs from Step 7 should be documented fully in the QAPP, sampling and analysis plan, or waste analysis plan. These outputs include the following (EPA 2006a):

1. Full documentation of the final sampling and analysis design, along with a discussion of the key assumptions underlying the design
2. Design implementation details and contingency plans for unexpected events
3. The QA and QC features that would be performed to detect and correct problems and ensure defensible results

Detailed descriptions of Step 7 activities are found in EPA 2006a (QA/G-4), pages 71–80.

5 Data Quality Assessment

The data quality assessment (DQA) phase naturally follows the DQO and field sampling stages (EPA 2006b). Whereas the sample planning stage for piles is anticipated to be more rigorous than many field sampling programs, the DQA phase is likely to be more streamlined due to the use of incremental sampling approaches. The DQA phase applies scientific and statistical methods to evaluate whether data achieve the type, quality, and quantity goals specified in the DQO stage. The DQA step fulfills the assessment phase in Figure 1 and involves five steps. A more detailed presentation of the DQA steps appears in Figure 18.

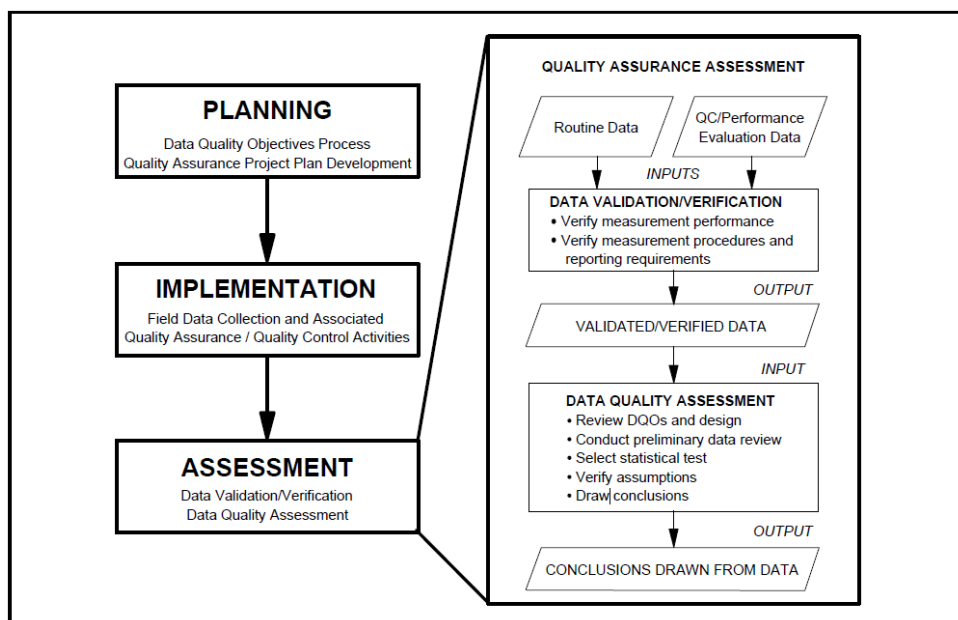


Figure 18. Data quality assessment in the quality process.

5.1 Step 1: Review the DQOs and Sampling Design

Step 1 ensures the DQOs established still apply to the project goals. Sample data collected during project implementation may challenge assumptions, conceptual models, or other DQO features. This is a QC step to verify the DQOs remain valid.

5.2 Step 2: Conduct Preliminary Data Review

Step 2 consists of reviewing QA reports, calculating basic statistics, and generating graphs of the data (if needed). If incremental sampling is used, the statistical and graphing tasks will be simplified. If data for estimation are collected, graphing (e.g., histograms, cumulative frequency plots, regression diagrams) and statistical analysis will be more extensive.

5.3 Step 3: Select the Statistical Test

Step 3 may also be simplified if using incremental sampling, where the mean of sample data is most applicable unless assumptions change. For estimation problems, this step needs to be performed in full.

5.4 Step 4: Verify the Assumptions of the Statistical Test

Step 4 evaluates the validity of assumptions based on the field data collected. If departures from original assumptions exist, step 4 determines whether the departures are acceptable.

5.5 Step 5: Draw Conclusions from the Data

Step 5 implements the chosen statistical test and draws conclusions regarding the status of the waste. The conclusions and inferences drawn from the statistical test should be documented fully.

Detailed guidance on the DQA process is provided in EPA 2000.

6 Incremental Sampling Methodology

The incremental sampling methodology (ISM) manages sampling errors through a structured approach to composite sampling, processing, and analytics. ISM increases the representativeness of samples and greatly reduces intersample variability, which provides additional confidence in decision making. ISM is highly useful for projects where the mean concentration of a constituent of interest needs to be established. Pile sampling is a situation where the mean concentration will be used for decision making; thus, ISM is useful for sampling and appraising soil piles.

Incremental sampling derives its name from the increments of material that are collected to represent a soil pile or other mass (Gy 1976; Pitard 1993; Myers 1997; Ramsey and Hewitt 2005). Sample increments are similar to discrete samples but have a predefined physical mass for sampling and laboratory subsampling; a defined procedure for field and laboratory processing (combination, drying, particle size reduction, subsampling, etc.); and a conceptual model of the soil material and contaminants in the soil. These attributes make ISM a structured and more sophisticated approach than composite sampling.

6.1 Heterogeneity and Its Impact on Sampling and Decision Making

Heterogeneity is the main source of variability and uncertainty in sampling (Pitard 1993; Myers 1997). Heterogeneity operates at several levels and in two ways, each of which influences sampling errors. It is important to understand the types of heterogeneity and their influences on sampling errors and decision making so the DQO process develops an appropriate sampling plan.

Figure 19 shows a conceptual model of variability and heterogeneity in particulate materials such as soils, sediments, and other potentially contaminant-bearing materials. It shows the influence heterogeneity has on sampling variability and accuracy of the sample results. The small dots represent the presence of a contaminant in a soil matrix. The small boxes represent small, discrete samples.

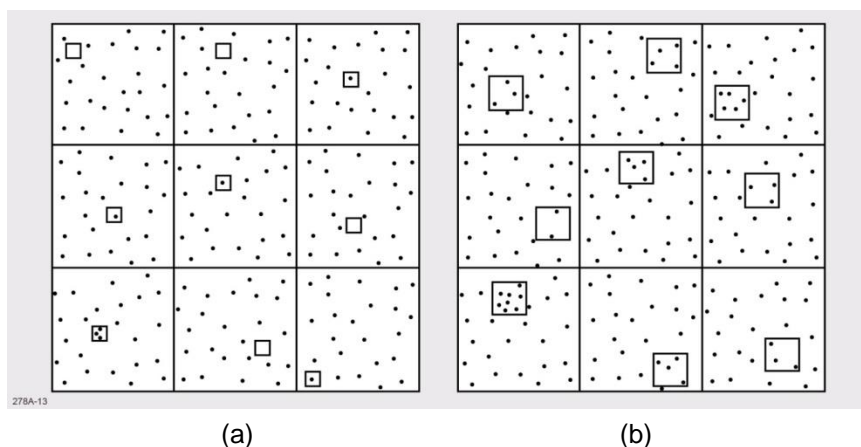


Figure 19. Conceptual model of heterogeneity and sample support.

While these dots are spaced fairly evenly, variability is present. Based on the sample support (size) and the contaminant density in Figure 19(a), one expects to capture or not capture the contaminant about 50% of the time. Note that the expected value for any given sample is 0.5 particles. Thus, finding the average would take at least two samples because in most cases the sample either misses the contamination or collects one particle, which produces a laboratory analysis that is either nondetect or twice the true average.

When a small amount of material is collected, either in the field or during laboratory subsampling, variability and bias problems occur. In Figure 19(a), approximately 44% of the time none of the contaminant is captured in the sample, and laboratory results come back as nondetects. Similarly, 44% of the time, one particle is captured and overestimates the mean. Occasionally, clusters are captured and over-represent the contaminant concentration (small box in lower left corner), causing potentially nonhazardous material to be classified as hazardous. In a process control situation, the cluster result usually generates one of the following conclusions: (1) the process is out of control or (2) the laboratory made an analytical error. Figure 19 shows neither of these conclusions to be true; the cause of the “spike” was sampling error.

Heterogeneity randomizes each particular result and thereby creates the need to collect a large number of samples to have a representative portion of the material for decision making. Large numbers of samples equate to large project costs; thus, ISM provides a valid mechanism to

increase decision-making quality while moderating expenses. Typical ISM projects for in situ soils collect 30 to 60 increments.

ISM recognizes the need to alter thinking and practice regarding the physical collection of samples by attributing importance to the amount of physical mass collected. Traditionally, the physical mass aspect has not been recognized as important to sampling. ISM demonstrates how the physical mass, or *support*, is a crucial component in characterization through sampling even if this aspect is nonintuitive. By using larger sample supports (Figure 19(b)), each sample captures some of the contaminant of interest, although not always at the correct proportion. More appropriate sample supports are shown in Figure 19(b).

Figure 20 shows how the support effect influences the statistical distribution of sample data. When small support samples are taken, often the most likely result is to find no contamination. Samples of insufficient support frequently produce skewed distributions, with most of the samples in the low concentration ranges. This implies the mode (most frequently observed value) is nondetectable for the contaminant of interest. As sample support increases, contaminant capture increases correspondingly and the mode shifts to the right. If sample support is sufficient, the mode becomes the central value and mimics a normal distribution, which is the most desirable result (and sample support). In this scenario, the most likely sample value equals the mean or a value close to the mean, which is the objective of ISM.

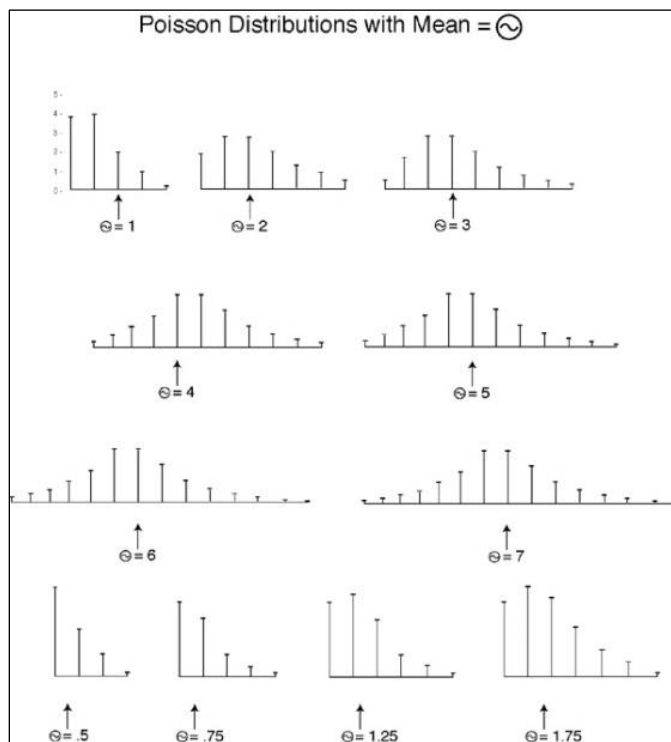


Figure 20. Statistical distribution variation based on sample support.

One of the unstated objectives of the DQO process is to “begin with the end in mind.” Figure 20 speaks to this point. If your objective is to analyze sample results using normal distribution statistics, you must use appropriate sample support for the samples collected.

Figure 21 shows an alternative conceptual model of contaminant distribution. In this case, the contamination occurs in larger particles and is more difficult to detect with small sample supports, which overly biases the results when contamination is found. This situation occurs frequently in contaminants such as PCBs, explosives, and other constituents that do not distribute evenly. Increasing the sample support addresses the issue. Moreover, ISM has the capability to quantify the appropriate mass of sample material to collect (section 6.4). However, each contaminant in each soil pile needs to be evaluated by means of a conceptual model so the appropriate sample supports can be determined.

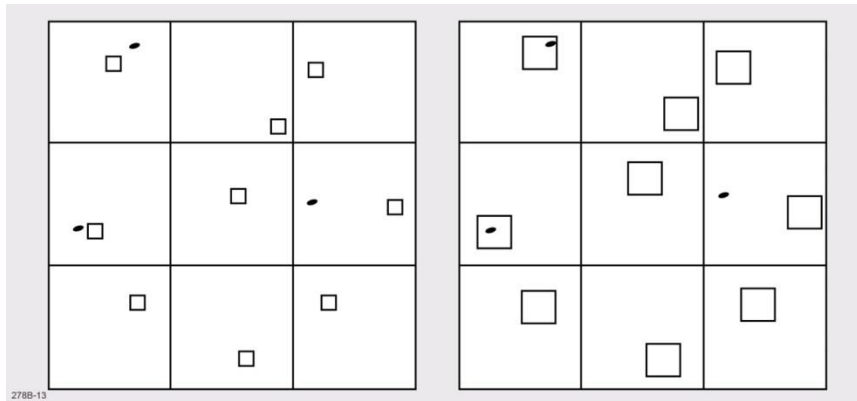


Figure 21. Conceptual model of heterogeneity: larger and dispersed contaminants.

6.2 Types of Heterogeneity

Three types of heterogeneity exist, each with different impacts on sampling design:

1. Constitution heterogeneity
2. Distribution heterogeneity
3. Zero-, one-, two-, and three-dimensional lots

The conceptual heterogeneity shown in Figure 19 and Figure 21 becomes more complicated in field situations. Heterogeneity in soils takes on two different forms, *constitution heterogeneity* and *distribution heterogeneity*. These two types of heterogeneity were first quantified by Pierre Gy (1976) in his STP. Each has its impact on sampling errors; thus, it is important to recognize the differences between these two heterogeneities, their influences on sampling errors, and how to mitigate their impacts. Constitution heterogeneity and distribution heterogeneity will be discussed; lot dimensions are beyond the scope of this paper.

6.2.1 Constitution Heterogeneity

Constitution heterogeneity refers to the fundamental properties of the soil material fragments when they are examined one by one. As a practical matter, this task is impossible. However, a quick examination of what appears to be homogeneous material, such as uniformly sized sand particles, reveals variations in size and shape. Constitution heterogeneity is the starting point for heterogeneity evaluation. It is a primary structural property of the material and cannot be altered by mixing or homogenization techniques. Only physical alteration by size reduction or another technique can alter the constitution heterogeneity.

Figure 22 shows a close-in view of soil particles. The figure shows the variety of sizes and shapes that make up the soil material. As seen in Figure 19 and Figure 21, taking small samples of this material could result in a wide variety of sample results.

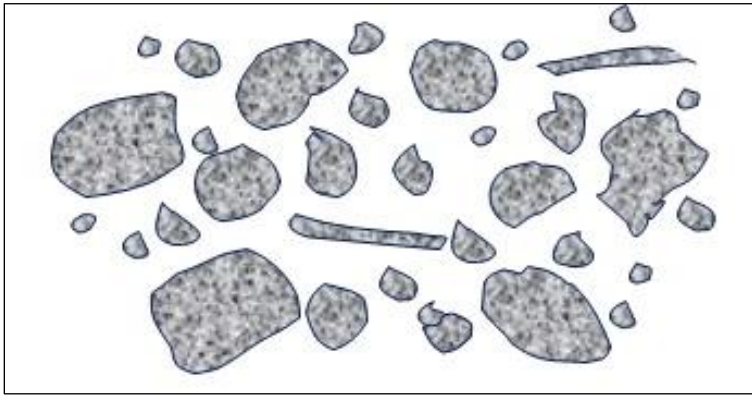


Figure 22. Constitution heterogeneity.

Constitution heterogeneity manifests itself by causing the fundamental error (FE), the first of Gy's sampling errors. The FE cannot be ignored, because once incurred, it never cancels out. It is a state of minimum error that represents the underlying variability of the soil or waste. As such, managing the FE is a priority. An extreme example of constitution heterogeneity is shown in Figure 23, which is a pile of auto fluff (ground-up automobile components).



Figure 23. Heterogeneity in auto fluff.

6.2.2 Distribution Heterogeneity

Distribution heterogeneity is both a complement to and a function of constitution heterogeneity; it is only possible if constitution heterogeneity is present. Distribution heterogeneity addresses the spatial component of particle distribution that is not addressed by constitution heterogeneity.

However, both types of heterogeneity are influenced and managed by appropriate sample support volumes.

Figure 24 shows the spatial nature of distribution heterogeneity. Note how the particles in the left oval are on average larger than those in the right oval. If these represented two laboratory subsamples, different analytical results could result from these two groupings.

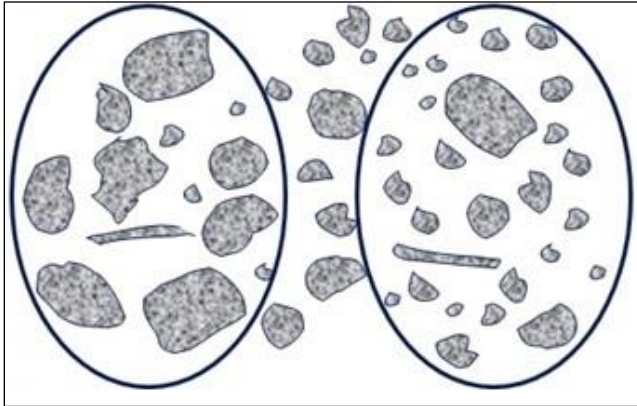


Figure 24. Distribution heterogeneity.

Distribution heterogeneity is responsible for *grouping and segregation error* (Figure 3). This error depends on three factors: (1) the constitution heterogeneity, (2) the spatial distribution of the constituent particles in soil, and (3) the shape of the pile. The pile shape is not trivial as gravity plays a key and never-ending role in the redistribution of particles.

Random particles of different natures tend to cluster naturally, including nonsoil materials. Clustering contributes heavily to the grouping factor no matter what the material. Figure 25(a) shows a mixture of peas and carrots where both constituents are approximately the same size, although differing in shape. This random mixture exhibits clustering of both items; they are not distributed evenly with a checkerboard-like pattern. No amount of mixing or homogenization will alleviate this condition, and small support samples taken from different locations will provide wide ranging results.



Figure 25. Heterogeneity and clustering with two constituents.

Figure 25(b) shows these two constituents where the carrots are more naturally shaped (long and thin). This example is a conceptualization of longer, flatter clay particles mixed with more spherical sand particles. The shape change from (a) to (b) complicates the heterogeneity further.

Figure 26 shows randomly dispersed gumballs with clusters of various colors, which conceptually is more like soil in that multiple types of particles are usually present. Again, these clusters do not suggest improper mixing or homogenization strategies. If the gumballs were mixed thoroughly again, new clusters would appear. Selecting a sufficient sample support is crucial in minimizing clustering effects. Taking a tablespoon-sized sample could easily report an absence of one or more colors. Gumballs of different sizes would increase heterogeneity further.



Figure 26. Natural clustering with multiple constituents.

Similarly, sorting due to gravity or other forces can segregate material. Mechanical handling of pile materials can also result in sorting and segregation. Other processes such as rain or wind can separate finer particles from larger particles and cause segregation, which in turn causes sampling errors. Again, appropriate sample supports mitigate the effects.

6.3 Sampling Correctness

Sampling strategies are sometimes chosen based on tradition, convenience, and cost. These can be desirable attributes, but they do not ensure proper sample results. Similarly, the goals of a sampling program are often precision and accuracy (Figure 27), which is also a misguided approach. The focus of a sampling program should be precision and *correctness* (Pitard 1993; Myers 1997), not precision and accuracy. If a sampling program is correct, it is likely to be accurate. The converse is not true.

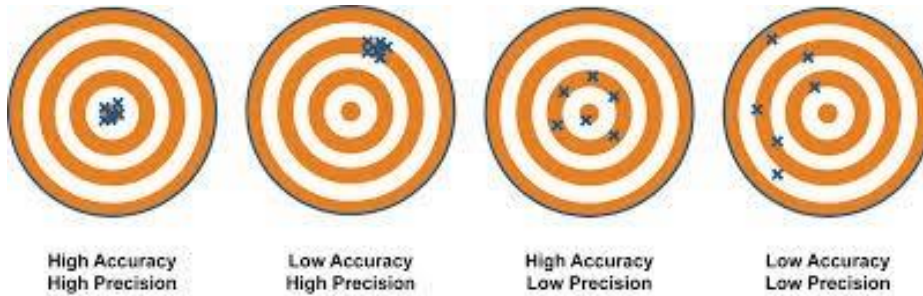


Figure 27. Traditional concept of accuracy and precision.

Just as constitution heterogeneity is a structural property of the pile material, sampling correctness is a structural property of the sampling strategy. This property is independent from the sampling problem itself. Correctness is intrinsic to the process and is a necessity if accurate results are desired. In other words, it is nonnegotiable if accurate data are desired. Whereas luck may provide accurate results, correctness ensures better and more accurate results. Accuracy is a circumstantial property and results from correct sampling processes.

Figure 28 is a matrix showing the relationships between correctness, precision, and accuracy. Sampling correctness is not influenced by uncontrollable circumstances such as pile material property. Fortunately, we may have control over the sampling process (correctness). As a circumstantial property, accuracy is influenced by (1) sampling correctness and (2) properties of the material. Correctness must also be present at each stage in the sampling and subsampling process. Failure to maintain correctness in subsampling introduces bias and inaccuracy, leading to decision errors.

Sampling Results	Sampling Structure	
	Correct	Incorrect
Accurate	Certain	Possible
Biased	Impossible	Probable

Figure 28. Results of correct and incorrect sampling.

Sampling correctness is linked primarily to the material heterogeneity and the sampling equipment used. Many common sampling devices are incorrect and bias the results in an unknown direction (high/low) and to an unknown extent (magnitude). Correct and incorrect sampling devices are discussed in section 6.4.

6.4 Sampling Errors

Heterogeneity is somewhat abstract; this section demonstrates how heterogeneity is linked directly to the following quantifiable sampling and decision-making errors.

1. Fundamental error
2. Grouping and segregation error
3. Long-range heterogeneity fluctuation error
4. Periodic heterogeneity fluctuation error

5. Increment delimitation error
6. Increment extraction error
7. Preparation error

6.4.1 Fundamental Error

The FE is the first sampling error because its source is the constitution heterogeneity, a primary structural property, which is fundamental to the material, hence the eponymous name. The FE possesses an insidious property in that it never cancels out once it has been made, even if the remaining steps in the sampling process are perfect. Fortunately, it is also the most quantifiable.

The magnitude of the FE is largely related to the sample support, with larger sample volumes producing lower FEs. Graphical illustrations of this concept appear in Figures 19 through 26. As shown in these diagrams, the error is made during sample conceptualization, even before the sample is taken.

The FE can be calculated using a constant factor of constitution heterogeneity C , the diameter of the largest particle in the material in centimeters (cm), and the mass of the sample in grams (Pitard 1993; Myers 1997). Ramsey (unpublished) has established the C value at 20. The formula for the fundamental error is

$$FE = \frac{Cd^3}{M}$$

where

FE = fundamental error (relative variance)

C = heterogeneity constant

d = diameter of the largest particles in centimeters

M = mass of the sample in grams

Another property of the FE is that each time a subsampling process occurs, a new FE is made. These sequential errors are additive. For example, if a bulk field sample is collected, a portion of the bulk material selected, and a laboratory subsample taken, three FEs are incurred. The initial FE is static; the size of subsequent FEs can be managed and reduced by comminution (crushing and grinding). Homogenization does not eliminate the FE.

Pitard (1993) provides a mechanism to manage the FE at each stage. Figure 29 is a nomograph that allows calculating and tracking of the FE incurred at each sampling stage. The graph tracks the mass (weight) of the sample along the x -axis, the FE on the y -axis, and the particle diameter size with the diagonal lines. The figure shows an initial 10-kilogram sample (A) with maximum particle diameters of 0.475 cm. Next, a subportion of material is removed (B). Point B falls on the horizontal line corresponding to the DQO-established maximum tolerable FE at each sampling stage, typically around 10–15%.

After the subportion is available, it undergoes size reduction to a maximum particle diameter of 0.170 cm and the FE goes to point C. This subsampling and comminution process continues until point H, where the sample is analyzed by the analytical laboratory equipment.

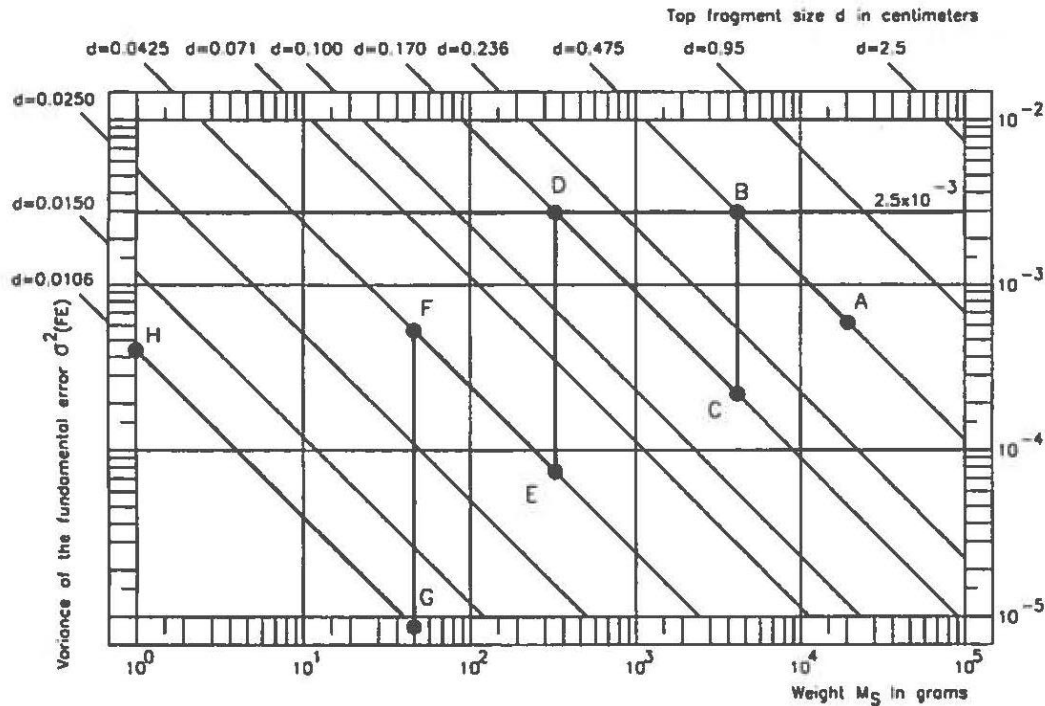


Figure 29. Nomograph for tracking the fundamental error (Myers 1997).

6.4.2 Grouping and Segregation Error

As shown in Figure 22 and Figure 24, soil particles group and segregate due to both randomness and sorting processes. Vibrations or shaking, such as during sample transport to the analytical laboratory, can cause smaller particles to move downward leaving larger particles near the top. Because clay and other smaller particles often contain higher concentrations of contaminants, a sample that was blended in the field may reach the laboratory in a segregated state. If the lab technician simply takes a small portion off the top of a jar, a biased sample results.

Blending, mixing, and other attempts at homogenization must be done with caution. Field methods such as coning and quartering are often ineffective and may exacerbate the segregation problem. Pitard (1993) calls vibratory spatulas “perfect segregating devices.” Using heavy machinery such as a backhoe is also not recommended for homogenization. Incremental sampling techniques offer a reliable approach to minimizing grouping and segregation error.

Riffle splitters offer mass reduction options at both the field and laboratory levels. Figure 30 presents a field riffle splitter; Figure 31 shows small-scale splitters for the laboratory. The splitter has dividers that channel soil material from a pan into two different collection bins. One of the bins is then randomly selected for further processing or analysis. A bulk sample can be sent through the riffle splitter more than once if necessary. Figure 32(A) shows that proper use of the splitter is important. If the pan is unequally loaded, more material may be collected in one pan (Figure 32(B)).



Figure 30. Field riffle splitter equipment.



Figure 31. Laboratory riffle splitters.

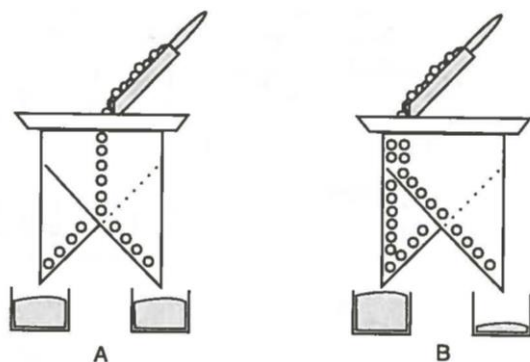


Figure 32. Proper (A) and improper (B) use of a riffle splitter.

6.4.3 Long-Range and Periodic Heterogeneities and Sampling Errors

These two errors are separate (Figure 3 and 40) but are discussed together because they are similar in nature and common. Samplers have long noted that in situ soils, sediments, and other solid materials often exhibit localized similarity. Thus, when samples are clustered near a single location, the concentrations are often similar. However, as one moves further away from the initial location, samples often exhibit increased variability. This phenomenon is termed *spatial correlation* (Myers 1997) and is common in environmental, mining, agricultural, and other applications where concentrations tend to vary in different portions of the site or at different depths. For various reasons, zonation may result in cyclical patterns over time and space.

Variability in concentrations over time and space creates concerns regarding the representativeness of a sampling program. Incremental sampling programs provide a high degree of areal coverage to increase representativeness of the average concentration.

6.4.4 Increment Delimitation and Extraction Errors

Increment delimitation and increment extraction errors are closely related and are discussed in tandem. The increment delimitation error challenges sampling technology and budgets. For three-dimensional materials, the ideal sample or sample increment is a sphere, which is usually impossible. For two-dimensional materials, the ideal increment is a vertical cylinder such as a core sample. Even cores prove problematic as recovery is often less than 100% or the extracted core includes material from under the pile.

Common sampling devices inherently fail to delimit the appropriate sample material. Figure 33 shows three such devices: a thief (A), an auger (B), and a trier (C). The thief has two obvious issues. The end portion is cone-shaped and excludes material from the bottom of the pile and the holes in the sides allow other material to fall out. The auger is even less contained and also excludes the bottom material. Moreover, it is likely to sample inappropriate, nonpile material from beneath the pile, biasing the sample. The trier exhibits the same problems as the thief.

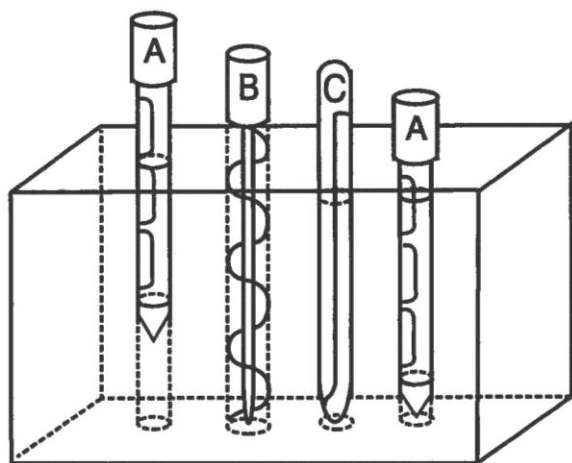


Figure 33. Thief, auger, and trier sampling devices.

Other common field and laboratory devices are also incorrect. Figure 34 shows that spatulas, scoops, and shovels are incorrect. Correct tools are designed to have flat bottoms and perpendicular sides of appropriate height.

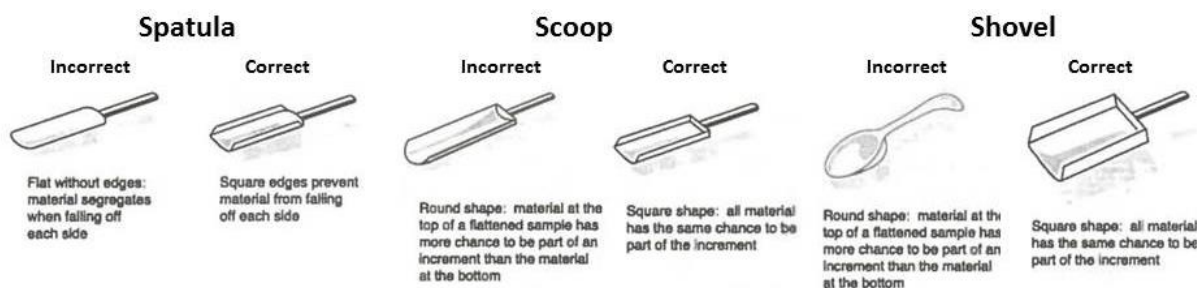


Figure 34. Incorrect and correct sampling devices.

Figure 35 shows the delimitation issue with spatulas (area A). The correct delimitation is shown on the right—a rectangular shape (area B). Because of gravity, collecting this material is not possible because of slumping, and some of the top portion of the material is excluded from the increment in both delimitation and extraction. Because of grouping and segregation, this biases the sample to collect less of the fine-grained material. The opposite problem occurs in Figure 36, where the scoop fails to collect all of the material on the bottom of the pile. These piles are called Japanese slab cakes (Pitard 1993) and can be constructed in the field or the laboratory and. Figure 37 shows how a slab sample is consistently biased by using incorrect tools such as spatulas or scoops.

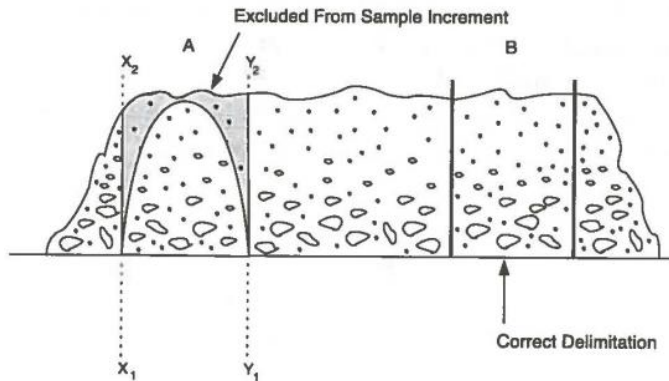


Figure 35. Increment delimitation and extraction error for spatulas.

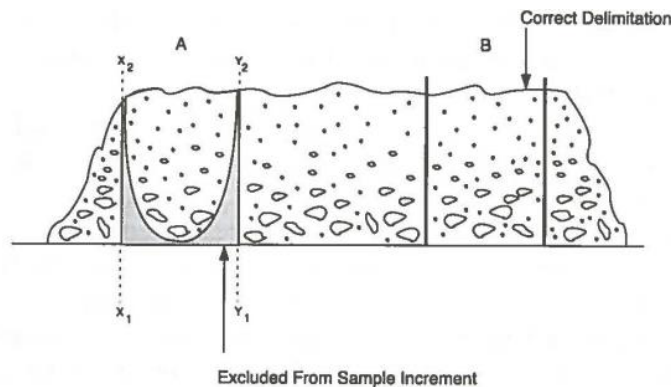


Figure 36. Increment delimitation and extraction error for scoops.

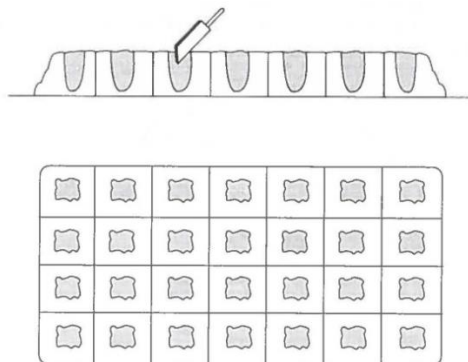


Figure 37. Increment delimitation and extraction error for spatulas or scoops on Japanese slab cakes.

A better method is to use a square-edged device to extract the entire cross-section of a linear pile of material (Figure 38). In theory, this technique could apply to operations such as windrow composting, but the practical aspects are limiting.

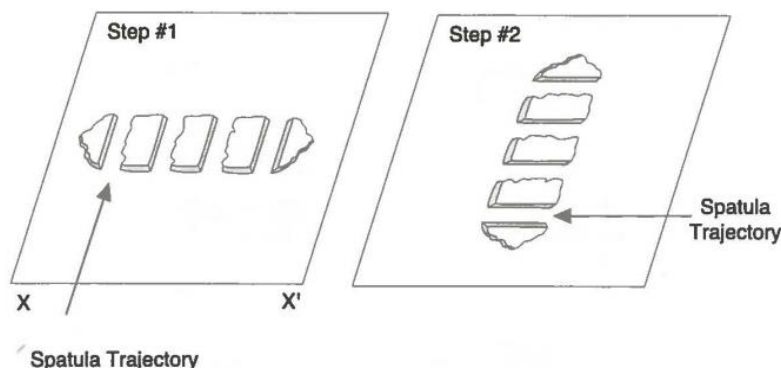


Figure 38. Correct application of the Japanese slab cake technique.

Piles are essentially large-scale, less uniformly shaped Japanese slab cakes. Increment delimitation and extraction (Figure 39 (a) and (b), respectively) are subject to the same sampling errors (Pitard 1993). Even if piles are flattened, edge effects present sampling challenges.

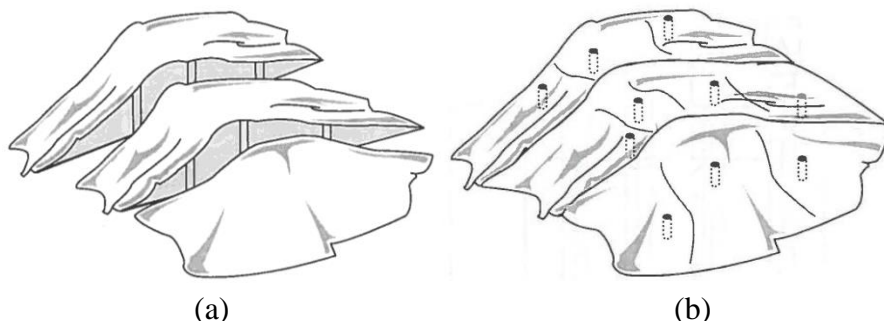


Figure 39. Increment delimitation (left) and extraction (right) in piles.

6.4.5 Preparation Errors

Preparation errors, which occur when material is rendered for analytical equipment, constitute a threat to quality due to the biases they introduce. These errors differ from other sampling errors in that they are not part of the selection processes described previously. They occur at several preparation stages and processes including the following (Myers 1997):

- Comminution stages, such as grinding, crushing, and pulverizing to reduce particle sizes
- Screening (wet or dry), which commonly accompanies comminution
- Drying solids to remove moisture
- Filtration to separate solids from liquids

Actual preparation errors include the following:

- Contamination errors, which introduce inappropriate material to the sample. In pile sampling, this can occur during increment extraction.
- Loss errors, which often occur during crushing and pulverizing operations. These can be significant for VOCs.

- Alteration of chemical composition, which is often associated with loss errors.
- Unintentional mistakes, typically resulting from operator error.

Whereas these errors impact quality, much attention is given to them in EPA and analytical laboratory guidance and they are only mentioned here. The reader is directed to Gerlach and Nocerino (2003) and other sources for details.

Figure 40 shows a detailed depiction of Gy's sampling errors, starting with the FE and ending with preparation error. The discrete model groups errors that occur essentially at a point in space; the continuous model groups errors that occur in two- or three-dimensional space. The diagram shows how preparation error occurs last and is part of neither the discrete or continuous model.

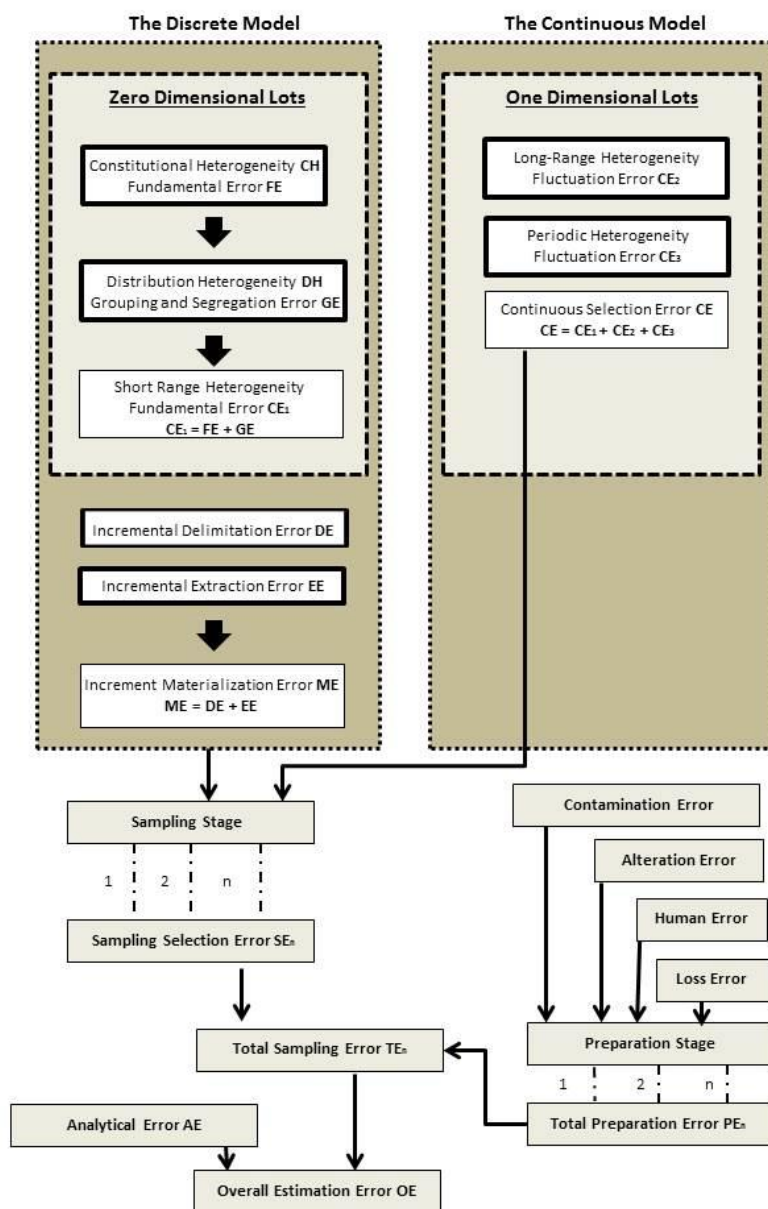


Figure 40. Detailed depiction of Gy's sampling errors.

6.5 Applying Incremental Sampling to Piles

This section focuses on the practical aspects of collecting incremental and/or other types of samples, including the following:

1. Field implementation, including pile access issues and increment support and number
2. Sampling tools
3. Field sampling
4. Laboratory processing and subsampling

The DQO process develops the goals and objectives for incremental sampling of piles. The sampling and analysis plan should contain the specific details regarding how the sampling will be conducted. At this point, the following items should be documented:

- Pile conceptual model
- Conceptual model of heterogeneity
- SUs and/or DUs
- Number of increments to compose a sample
- Number of incremental samples to be created and analyzed
- Mass of material in an increment
- Locations/depths for increment collection
- Sampling tool types
- Sample extraction protocols
- Field increment sampling and processing procedures (sieving, drying, volume reduction, etc.)
- Laboratory subsampling and processing procedures
- Analytical methods

The overall objective is to maintain sample quality so decisions can be made with confidence. The DQO process has provided a path to minimize errors on multiple levels; field implementation must be executed with precision to maintain the planned quality for subsequent decisions.

6.5.1 Pile Access Issues

Piles are sampled most easily when they are created. If truckloads or backhoe scoops are available for sampling, increments can be taken at this time assuming the number of loads or scoops exceeds the number of increments required. For existing piles, a more common situation, flattening the pile to a manageable depth can ease and facilitate sampling. A depth of 3 feet or less is recommended (ITRC 2012). Figure 41 shows a flattened pile with a grid for incremental sampling marked off. Note that the pile in Figure 41 is relatively small, implying the SU and DU may be the same. Larger piles may be partitioned into appropriately sized areas for decision making.



Figure 41. Flattened pile with incremental sampling grid (ITRC 2012).

Large piles or processing piles are more difficult to access. Figure 42 presents a conceptual windrow processing pile and its dimensions. Operational constraints often limit the height and width of the pile, creating access challenges. Due to the angle of repose, only a narrow area is available on the top of the pile, and placing people and sampling equipment on the narrow ridge is challenging. Similarly, a large interior and base exist, both of which will be difficult to delineate and sample correctly with common sampling tools.

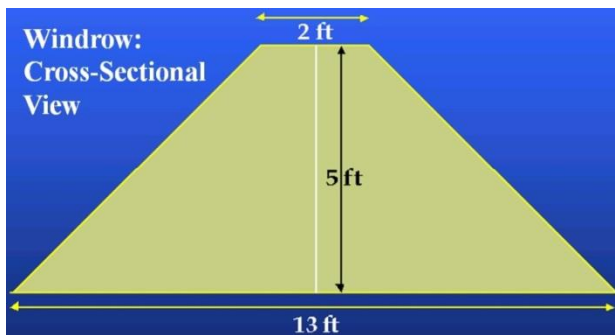


Figure 42. Cross-sectional view of a windrow composting pile.

Figure 43 shows a plan view with proposed increment collection locations. Physically, these samples may be difficult to collect for access and stability reasons. Moreover, extracting a full vertical profile of the material is unlikely. The issues are similar for large, thick piles; however, a broader top may allow equipment access for coring or other cylindrical device sampling.

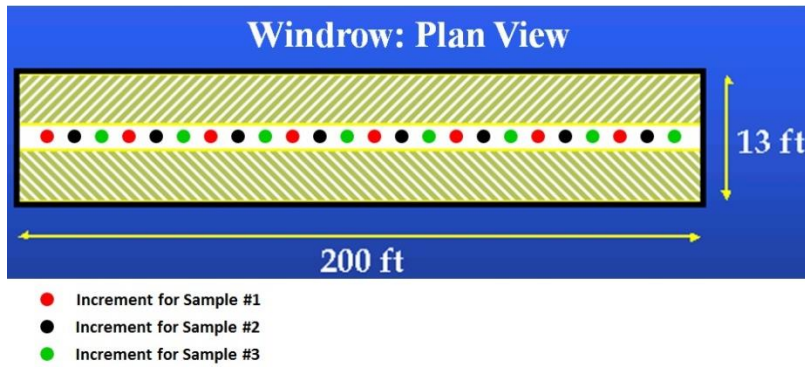


Figure 43. Plan view of a windrow composting pile.

Figure 42 illustrates how pile edges will always be an issue. From a sampling representativeness perspective, these areas should not be excluded, otherwise the sample is biased. The edges, especially in outdoor piles, are subject to weathering and the material particle sizes may be segregated more than the bulk, interior areas of the pile. This potential grouping and segregation should be incorporated into the conceptual model and sampling plan to minimize bias.

Figure 44 shows a sampling plan for a large pile where three incremental samples will be collected and analyzed. The red circles indicate increment collection locations for sample one, gray × symbols are for sample two, and blue triangles are for sample three. The switchback lines propose sample collection paths for the three samples. Figure 45 shows a similar model for large, round piles. Note how small, round piles pose challenges similar to windrow composting piles.

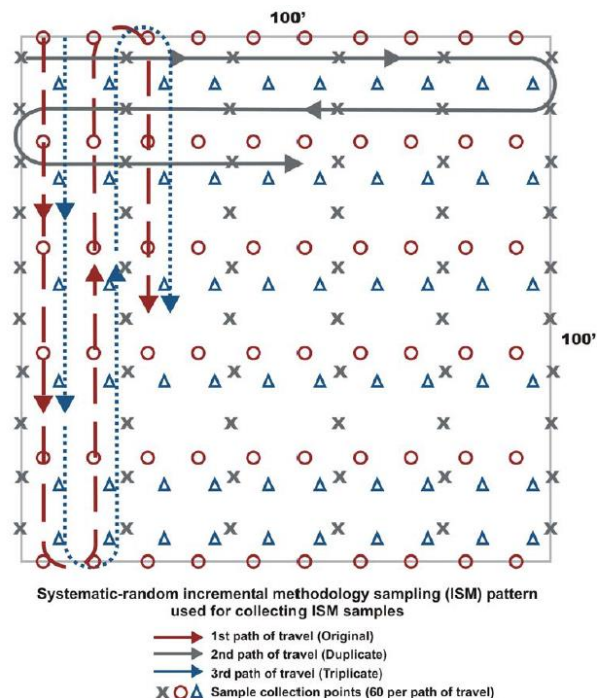


Figure 44. Increment collection grid for a large, rectangular pile.

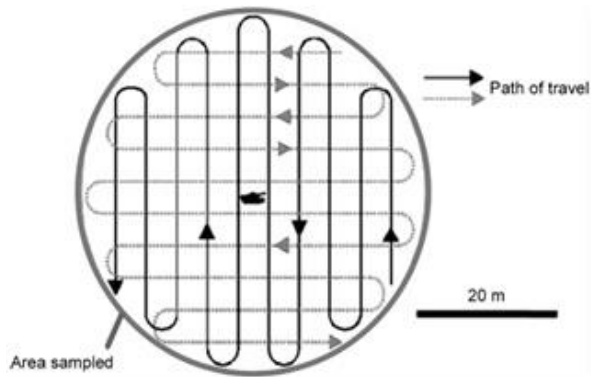


Figure 45. Increment collection grid for a large, round pile.

Figure 44 and Figure 45 show fixed grids for sampling. These are simpler to implement than random sampling as less effort is required to spatially locate random points to precise x,y coordinates. When implementing a fixed grid, it is recommended to randomly orient the grid origin to mitigate bias issues. The location of the sample increments will be at a fixed location within the grid cell. This preselection approach minimizes field sampler temptation to select the increment collection location in the grid based on visual appeal, which can bias results.

6.5.2 Increment Support and Number

The increment support (size) and number should reflect the conceptual model of heterogeneity, the maximum particle size, the minimum mass (weight) of sample material to satisfy the FE DQOs, and the practical field constraints. The FE is commonly limited to between 10 and 15% per sampling and subsampling stage (Pitard 1993; Myers 1997) and can be calculated using a nomograph (Figure 29). The conceptual model of heterogeneity (Figure 19 and Figure 21) influences the size of individual sample increments. Bulk samples may be required in order to obtain sufficient material to satisfy the FE constraints.

Pile accessibility influences the increment support as well. Core samples to access remote portions of the pile may collect more mass per increment than required. Field subsampling may be required to limit the amount of material that contributes to the incremental sample. Similarly, cores may be needed if large particles are anticipated to be encountered during increment extraction. Optimally, the core diameter should be larger than the largest particle, but in practice this is not always possible. Figure 46 shows how a core sample can collect particles only up to a given size. Conversely, if maximum particle sizes are relatively small, smaller mass increments may suffice. Increments for in situ soil sampling weighing between 20 and 60 grams are common (ITRC 2012).

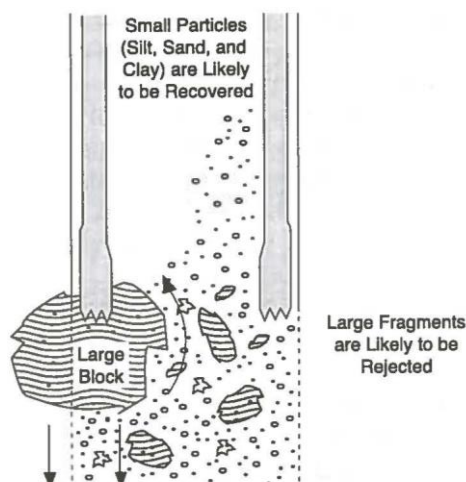


Figure 46. Core sampling issues.

In general, a minimum of 30 increments should compose an incremental sample. Public literature reveals situations where up to 100 samples were collected due to relatively large particle sizes in the constituents of interest (e.g., explosives, particulate metals, PCBs). The size and spatial variability of contamination may also influence the number of increments taken. As a rule of thumb, 30 to 60 sample increments provide a reasonable amount of coverage. To put this into a statistical context, the SER approach (EPA 2002a) provides a 95% confidence that 90% of the distribution concentrations have been captured when 30 increments are collected. Thus, one expects to collect increments at or close to both the highest and lowest concentrations present with a known level of confidence. If 60 samples are taken, the SER yields a 95% confidence that 95% of the distribution concentrations have been captured. With 90 samples, the results are 99/95% confidence.

Typical ISM field samples for in situ soils weigh 500–2,500 grams (1.1–5.5 pounds). For numerous reasons, substantially more material may be obtained during increment collection. The project DQOs may limit particle size, which may require field sieving to remove large particles and unwanted organic materials (Figure 47). This process simultaneously reduces the mass. Further mass reduction may be accomplished through the use of a riffle splitter. A common but incorrect way to reduce mass is via the Japanese slab cake method.



Figure 47. Soil sieving in the field.

The incremental sampling approach produces a single sample of multiple increments intended to represent the mean (average) concentration of a COC. Regulatory guidance or protocols commonly call for a UCL on the mean to meet DQO decision quality requirements. Calculating a UCL requires a sample standard deviation, which requires at least three samples. The three-sample minimum applies whether the samples are discrete, composite, or incremental.

6.5.3 Sampling Tools

A variety of sampling tools are available for collecting incremental samples. Those that are most common and easy to apply generally are incorrect in terms of increment delineation and extraction, as discussed in section 6.4. These include thief samplers, augers, triers, scoops, shovels, and spatulas. Core samples provide better delineation but can be difficult to deploy for the reasons stated previously.

As a practical matter, the cohesiveness and composition of the soil matrix will also dictate the types of sampling tools applied. Loosely consolidated materials offer the opportunity to reach remote portions of the pile material, but it may not be possible to recover all of the material upon extraction. Specially designed sampling devices, preferably with cylindrical shapes, may be designed and applied if required. The diameters should also be at least three times the diameter of the largest particles.

Sampling tools do not need to be decontaminated while sampling within a DU; however, when transitioning to a new DU, decontamination is necessary.

6.5.4 Field Sampling Summary

Field sampling to create incremental samples involves numerous steps and decisions. Figure 48 (ITRC 2012) provides a conceptualization of the process flow. Note that decision point one is a key discriminator. Projects where VOCs must be analyzed follow two separate paths, which increases the level of effort.

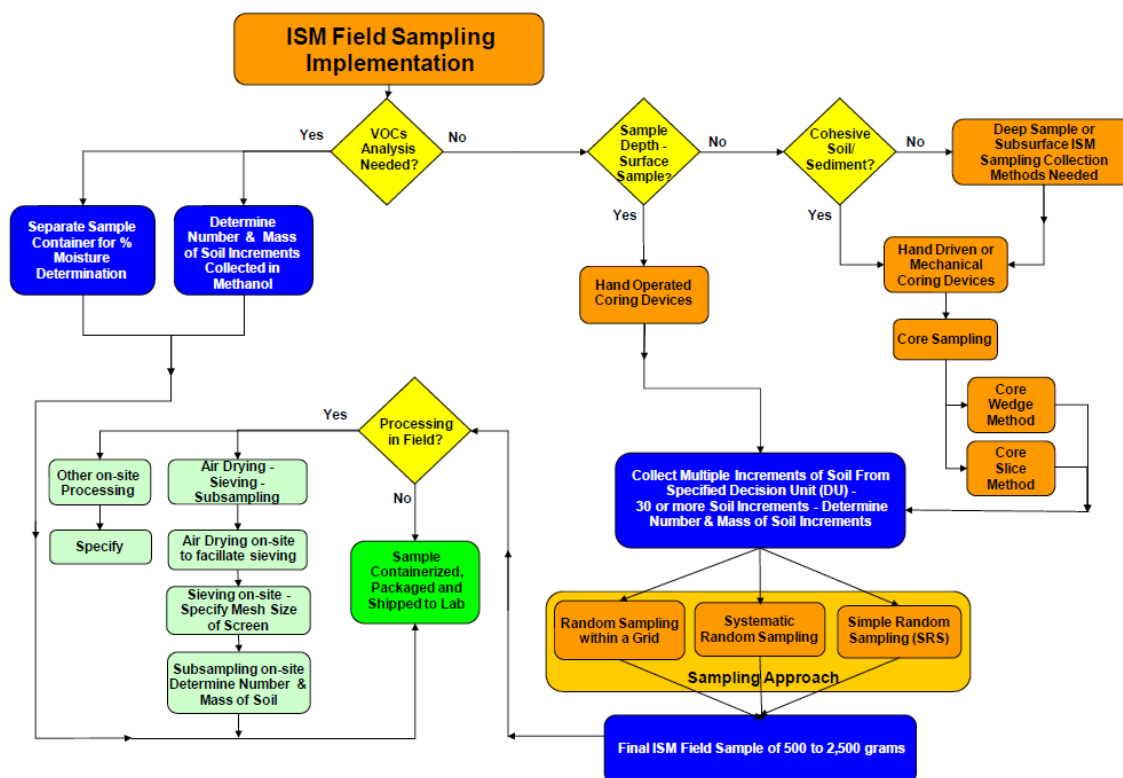


Figure 48. Flowchart of incremental sampling methodology.

6.5.5 Laboratory Processing and Subsampling

Because the FE increases at each subsampling stage, specific laboratory processing and subsampling procedures should be established in the DQO process and confirmed with the laboratory before samples are sent. Drying the sample, using air or other methods, is usually necessary before sieving or particle size reduction methods are applied. Particle size goals should be established during the DQO process. In some cases, sieving can dilute contaminant concentrations such as lead paint chips or ammunition fragments. The reverse also applies: sieving may enrich concentrations by eliminating larger, less contaminated particles. Note that samples scheduled for VOC analysis should not undergo sieving, drying, mixing, or particle size reduction as preparation errors will result.

If dried samples have formed clumps, sample disaggregation may be necessary to break up the particles. If the sample particles are sufficiently small and similar in size and density, mixing may be appropriate before subsampling. Caution is necessary as mixing and other homogenization attempts may increase grouping and segregation error. Size reduction is effective for reducing the grouping and segregation error; homogenization is not.

To reduce the particle size of solid materials before subsampling, mortar and pestle, rotary pulverizers, ball mills, and dish-and-puck (shatterbox) methods may be employed. Except for the mortar and pestle, these methodologies can reduce particles to $<100\text{ }\mu\text{m}$, with dish-and-puck methods capable of reducing particles to $<75\text{ }\mu\text{m}$. Caution should be exercised when grinding nitrocellulose-based propellant residues so that volatile energetic contaminants are not lost.

Grinding is not generally recommended for other organic constituents or other compounds susceptible to oxidation.

Once the laboratory has prepared the material received, splitting and subsampling processes must be applied as only a portion of the material will be sent to the analytical device. As with other steps, these procedures should be established in the DQO systematic planning process.

Depending on the DQOs, the following methods may be applied:

- Sectorial sample splitting using a rotary splitter. This method is effective in minimizing errors resulting from heterogeneities.
- Paper cone sectorial splitting. This method is an inexpensive *ad hoc* procedure but requires substantially more labor than the rotary splitter.
- One- or two-dimensional Japanese slab cakes (Figure 37 and Figure 38). This laboratory method is the same as for the field except using smaller volumes of material. Square-edged scoops are the correct sampling devices, not spatulas, spoons, or rounded scoops.

To create the subsample, at least 30 laboratory increments are recommended. Details regarding laboratory subsampling procedures can be found in Gerlach and Nocerino (2003).

7 Statistical Evaluation of Sampling Results

The initial presumption is that most decision statistics will be either average-based or presence/absence based. Discussion will focus in these areas:

1. Mean (discrete or incremental samples)
2. UCL on mean (discrete or incremental samples)
3. Presence/absence (discrete samples)

7.1 Incremental Samples

ISM inherently produce an estimate of the mean concentration of the constituents of interest. By combining material from portions of an entire DU and blending it before analysis, the analytical result represents the average concentration. Because incremental sampling uses 30 or more sample increments, in some situations a single incremental sample is sufficient for decision making.

Because of sampling and analytical errors, the analytical result may not represent the true mean concentration exactly. Whereas the DQO and laboratory processes have endeavored to minimize bias, a bias of unknown magnitude and direction may exist. For regulatory decision making, using a UCL on the mean provides additional assurance that hazardous material is disposed properly.

The standard error of the mean is calculated by the formula:

$$\sigma_M = \frac{\sigma}{\sqrt{n}}$$

where

σ_M = standard error of the mean

σ = sample standard deviation

n = number of samples

To obtain a standard deviation, a minimum of three samples must be available. Therefore, project DQOs should specify whether a single sample suffices or whether three ISM sample are needed so a UCL can be calculated. Once the UCL is available, it can be compared to the action level established in the DQO decision statement.

7.2 Discrete Samples

It is not mandatory to use incremental samples to characterize contaminant concentrations in piles. Discrete samples may also be used. For example, if a waste needs to be tested for a Resource Conservation and Recovery Act (RCRA) listed waste, discrete samples can be used to “search” for the presence of the COC. If incremental sampling is used and only a few increments capture the listed waste, the other increments may “dilute” the concentration below detection limits.

For various reasons, including the one mentioned above, certain types of RCRA sampling were required to use discrete samples instead of composite samples. Incremental sampling is now able to replace discrete sampling in some cases, although regulatory acceptance in all cases is not complete.

Accurate characterization of piles generally requires more than the one to three samples used in ISM. To meet DQO decision error tolerances, a precise number of samples required must be determined before field sampling begins. Once the samples have been taken and analyzed, the results are checked against DQO tolerances in the DQA phase to ensure the goals were met.

The decision performance goal diagram (Figure 16) can be used to estimate the number of samples required to meet the DQOs. If no sample data are available from the pile, an estimate of the standard deviation can be made. A conservative value is recommended to maximize the probability of passing the DQA review. However, the more conservative value, the more samples will be required. The DQO process is designed to evaluate costs, schedules, and available resources to optimize the estimates.

7.3 Presence or Absence

Determining presence or absence in a waste is often as important as knowing the average concentration. As stated previously, RCRA wastes must demonstrate they do not exhibit the characteristics of ignitability, corrosivity, reactivity, or toxicity. One common approach to determine presence or absence at a specified confidence level is the SER, which either requires that zero samples exceed the standard or allows for limited exceedances (EPA 2002a). A RCRA characteristic test may need to demonstrate zero samples exhibit the characteristics of concern, whereas other applications may allow detectable concentrations but with zero occurrences exceeding an established threshold.

Table 3 is an SER table that pairs the desired confidence level alpha (α) with the percentage (p) of the population covered. If the DQOs require a 90% confidence that 90% of the population is acceptable, 22 samples are needed. If 95%/95% is the goal, 59 samples are necessary.

Table 3. Sample size required to demonstrate with at least $(1 - \alpha)\%$ confidence that at least $p\%$ of a lot or batch of waste complies with the applicable standard (no samples exceeding the standard).

p	$1 - \alpha$										
	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	0.99
0.50	1	2	2	2	2	2	3	3	4	5	7
0.55	2	2	2	2	3	3	3	4	4	6	8
0.60	2	2	2	3	3	3	4	4	5	6	10
0.65	2	2	3	3	3	4	4	5	6	7	11
0.70	2	3	3	3	4	4	5	6	7	9	13
0.75	3	3	4	4	5	5	6	7	9	11	17
0.80	4	4	5	5	6	7	8	9	11	14	21
0.85	5	5	6	7	8	9	10	12	15	19	29
0.90	7	8	9	10	12	14	16	19	22	29	44
0.95	14	16	18	21	24	28	32	37	45	59	90
0.99	69	80	92	105	120	138	161	189	230	299	459

If one occurrence is tolerable, Table 4 may be applied. This approach is useful because random selection may by chance select an exceedance early in the sampling sequence, potentially disqualifying the pile from compliance even though the entire pile is actually compliant. Such a disqualification is a Type II error, where a compliant waste is judged to be noncompliant. By taking additional samples (Table 4), an early exceedance may be tolerated if no additional exceedances are observed. Note that the 90/90 and 95/95 confidence requirements go from 22 and 59 samples to 38 and 93 samples, respectively.

Table 4. Sample size required to demonstrate with at least $(1 - \alpha)\%$ confidence that at least $p\%$ of a lot or batch of waste complies with the applicable standard (one sample exceeding standard).

p	$1 - \alpha$										
	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	0.99
0.50	3	4	4	4	5	5	5	6	7	8	11
0.55	4	4	4	5	5	6	6	7	8	9	12
0.60	4	5	5	5	6	6	7	8	9	10	14
0.65	5	5	6	6	7	7	8	9	10	12	16
0.70	6	6	7	7	8	9	9	10	12	14	20
0.75	7	7	8	9	9	10	11	13	15	18	24
0.80	9	9	10	11	12	13	14	16	18	22	31
0.85	11	12	13	15	16	18	19	22	25	30	42
0.90	17	19	20	22	24	27	29	33	38	46	64
0.95	34	37	40	44	49	53	59	67	77	93	130
0.99	168	184	202	222	244	269	299	337	388	473	662

The SER is based on the hypergeometric distribution and has no distributional assumptions, which means the population can be normal, lognormal, or other distribution for a given sample support. This attribute allows flexibility and ease of use. However, the SER has limited power for small sample sets.

8 Process Flow

Pile sampling, characterization, and decision making require numerous information inputs and choices. Because of the complex nature of the task, no two pile sampling sequences will be identical. However, the overall process is consistent from one project to the next. It begins with systematic planning by implementing the DQO process. During DQO development, a decision will be made to use either incremental sampling or another method. Once this decision is made, error tolerances are established and a field sampling plan developed. After field implementation, laboratory analysis will provide the concentrations of the field samples—discrete, composite, or incremental. Once the data have been validated, a DQA is performed to ensure the DQOs were met. If not, additional field sampling may be necessary. If the DQOs were achieved, the waste may be classified for treatment and/or disposal. Figure 49 provides a flow diagram of the process.

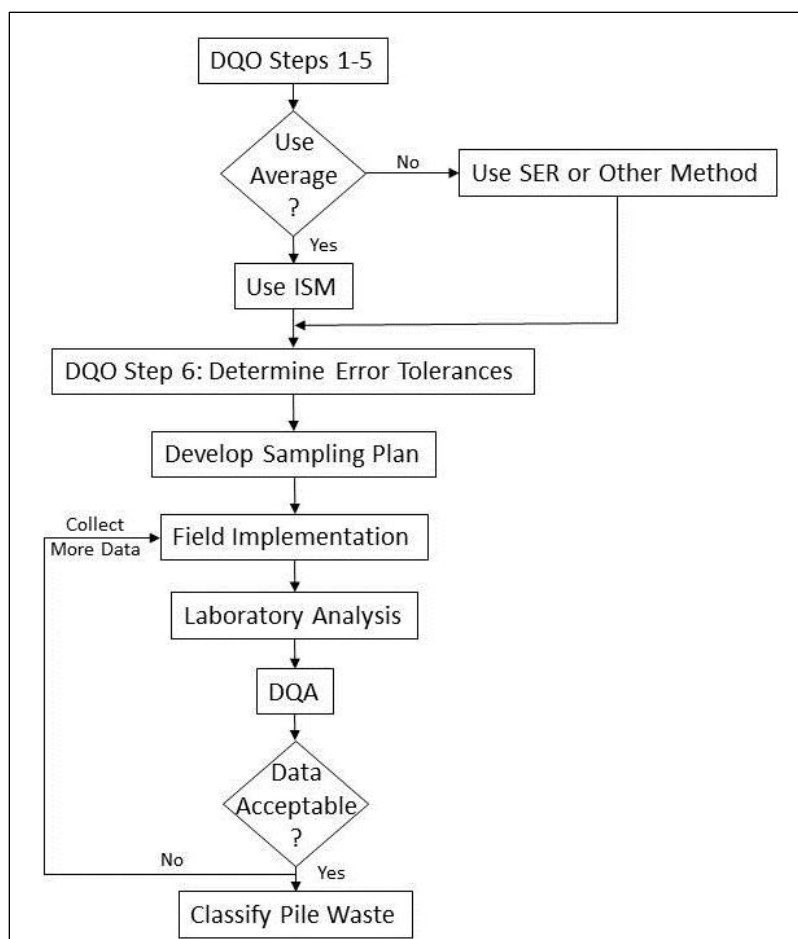


Figure 49. Pile sampling process flow.

9 Conclusion

Pile sampling represents a significant challenge to a quality system. The nature of the material, the logistics of obtaining correct samples, and limited time and budgetary resources combine to pressure the obtainable quality of the characterization. Pile sampling starts from a position of knowing the limitations of obtaining an accurate assessment. However, the need to reduce risk to human health and the environment through proper treatment and/or disposal of waste exceeds the concerns relating to characterization. The approaches and techniques described in this paper, if applied properly, can produce pile assessments acceptable to generators, regulators, and other stakeholders.

References

- EPA (US Environmental Protection Agency). 1995. *EPA Observational Economy Series, Volume 1: Composite Sampling*. Washington, DC: Office of Policy Planning and Evaluation. EPA-230-R-95-005.
<http://www.triadcentral.org/ref/ref/documents/composite.pdf>.
- EPA (US Environmental Protection Agency). 2000. *Guidance for Data Quality Assessment: Practical Methods for Data Analysis*. Washington, DC: Office of Environmental Information. EPA/600/R-96/084.
<http://www.epa.gov/region6/qa/qadevtools/mod4references/secondaryguidance/g9-final.pdf>.
- EPA (US Environmental Protection Agency). 2001a. *EPA Requirements for Quality Management Plans*. Washington, DC: Office of Environmental Information. EPA QA/R-2. <http://www.epa.gov/QUALITY/qs-docs/r2-final.pdf>.
- EPA (US Environmental Protection Agency). 2001b. *EPA Requirements for Quality Assurance Project Plans*. Washington, DC: Office of Environmental Information. EPA QA/R-5. <http://www.epa.gov/QUALITY/qs-docs/r5-final.pdf>.
- EPA (US Environmental Protection Agency). 2002a. *RCRA Waste Sampling Draft Guidance – Planning, Implementation, and Assessment*. Washington, D.C. Office of Solid Waste and Emergency Response. EPA 530-D-02-002.
<http://www.epa.gov/osw/hazard/testmethods/sw846/pdfs/rwsdtg.pdf>.
- EPA (US Environmental Protection Agency). 2002b. *Guidance for Choosing a Sampling Design for Environmental Data Collection*. Washington, DC: Office of Environmental Information. EPA QA/G-5S. <http://www.epa.gov/QUALITY/qs-docs/g5s-final.pdf>.
- EPA (US Environmental Protection Agency). 2006a. *Guidance on Systematic Planning Using the Data Quality Objectives Process*. Washington, DC: Office of Environmental Information. EPA QA/G-4. <http://www.epa.gov/QUALITY/qs-docs/g4-final.pdf>.
- EPA (US Environmental Protection Agency). 2006b. *Data Quality Assessment: Statistical Methods for Practitioners*. Washington, DC: Office of Environmental Information. EPA/240/B-06/003. <http://www.epa.gov/QUALITY/qs-docs/g9s-final.pdf>.
- Gerlach, R.W. and J.M. Nocerino. 2003. *Guidance for Obtaining Representative Laboratory Analytical Subsamples from Particulate Laboratory Samples*. EPA/600/R-03/027.
https://clu-in.org/download/char/epa_subsampling_guidance.pdf
- Gy, P.M. 1976. *Sampling of Particulate Materials*. Amsterdam: Elsevier.
- ITRC (Interstate Technology and Regulatory Council). 2012. *Incremental Sampling Methodology*. Washington, DC: ITRC. http://www.itrcweb.org/ism-1/pdfs/ISM-1_021512_Final.pdf

- Myers, J.C. 1997. *Geostatistical Error Management: Quantifying Uncertainty for Environmental Sampling and Mapping*. New York: John Wiley & Sons.
- Pitard, F.F. 1993. *Pierre Gy's Sampling Theory and Sampling Practice: Heterogeneity, Sampling Correctness, and Statistical Process Control*, 2nd ed. Boca Raton, FL: CRC Press.
- Ramsey, C.A. and A.D. Hewitt. 2005. "A Methodology for Assessing Sample Representativeness." *Environmental Forensics* 6:71–75.